

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2005, 15:29:19 ; Search time 12873 Seconds
(without alignments)
11352.515 Million cell updates/sec

Title: US-10-751-612-1
Perfect score: 3016
Sequence: 1 tctagacataggcattgta.....gggctatgtcaagtcctatgg 3016

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*
2: gb_hg.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	133.2	4.4	279448	8	AY661656 Sorghum b
C 2	122.8	4.1	268433	8	AV144442 Sorghum b
3	122.2	4.1	1215	8	L21752 Saccharum h
C 4	118.6	3.9	3688	6	AR412080 Sequence
C 5	118.6	3.9	3688	6	AR473380 Sequence
C 6	118.6	3.9	3688	6	AR488131 Sequence
C 7	118.6	3.9	3691	6	AR412075 Sequence
C 8	118.6	3.9	3691	6	AR473375 Sequence
C 9	118.6	3.9	3691	6	AR488126 Sequence
C 10	118.6	3.9	5174	6	AR412079 Sequence
C 11	118.6	3.9	5174	6	AR473379 Sequence
C 12	118.6	3.9	5174	6	AR488130 Sequence
C 13	118.6	3.9	5174	8	AF093505 Saccharum
C 14	106.4	3.5	183990	8	AF114171 Sorghum b
C 15	74.8	2.5	183990	8	AF114171 Sorghum b
C 16	73.8	2.4	157237	8	AY542311 Sorghum b
C 17	72.4	2.4	147461	8	AP004691 Oryza sat
C 18	72.4	2.4	148544	8	AP004636 Oryza sat
C 19	70.2	2.3	202197	8	AF466200 Sorghum b

C 20	69.8	2.3	136068	8	AC137693	AC137693 Oryza sat
C 21	68.4	2.3	3039	8	ZMU34726	U34726 Zea mays su
C 22	67	2.2	4636	8	AY177889	AY177889 Sorghum b
C 23	67	2.2	138858	8	AP002968	AP002968 Oryza sat
C 24	67	2.2	142376	8	AF503433	AF503433 Sorghum b
C 25	67	2.2	202197	8	AF466200	AF466200 Sorghum b
C 26	66.8	2.2	165745	8	OSJN00105	OSJN00105 Sorghum b
C 27	66.2	2.2	144159	8	AF466199	AF466199 Sorghum b
C 28	65.4	2.2	1334	6	AX406684	AX406684 Sequence
C 29	65.4	2.2	120132	8	AC120496	AC120496 Genomic s
C 30	65.4	2.2	159802	8	AP005447	AP005447 Oryza sat
C 31	65.4	2.2	162249	8	AF061282	AF061282 Sorghum b
C 32	64.8	2.1	121894	8	AC135599	AC135599 Oryza sat
C 33	64.4	2.1	144159	8	AF466199	AF466199 Sorghum b
C 34	64.4	2.1	150286	2	OSJN00241	OSJN00241 Sorghum b
C 35	63.8	2.1	135991	8	AP002882	AP002882 Oryza sat
C 36	63.8	2.1	162249	8	AF061282	AF061282 Sorghum b
C 37	63.8	2.1	182779	8	AP002845	AP002845 Oryza sat
C 38	63.4	2.1	165909	8	AP005420	AP005420 Oryza sat
C 39	63.4	2.1	187410	8	AP005579	AP005579 Oryza sat
C 40	62.8	2.1	103167	8	AY661658	AY661658 Sorghum b
C 41	62.8	2.1	125927	8	AY661657	AY661657 Sorghum b
C 42	62.6	2.1	795	11	AF344051	AF344051 Oryza sat
C 43	62.4	2.1	113332	8	CNS08C72	AL731750 Oryza sat
C 44	62.4	2.1	151117	8	CNS08C96	AL732531 Oryza sat
C 45	61.8	2.0	148736	8	OSJN00134	AL662944 Oryza sat

ALIGNMENTS

RESULT 1
AY661656
LOCUS AY661656 279448 bp DNA linear PLN 03-JUL-2004
DEFINITION Sorghum bicolor clone BAC 88M4, complete sequence.
ACCESSION AY661656
VERSION AY661656.1 GI:49359135
KEYWORDS HTG.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor

REFERENCE 1 (bases 1 to 279448)
Islam-Faridi, M.N., Kim, J.-S., Klein, P.E., Stelly, D.M., Price, H.J., Klein, R.R. and Mullet, J.E.
AUTHORS Cytogetic Analysis of Sorghum Chromosome 3 and Alignment to Rice Chromosome 1 Reveals Expansion of Pericentromeric Heterochromatin in Sorghum
JOURNAL Unpublished (2004)
REFERENCE 2 (bases 1 to 279448)
Klein, R.R., Klein, P.E., Mullet, J.E., Minx, P. and Miner, T.L.
AUTHORS Direct Submission
TITLE Submitted (21-JUN-2004) USDA-ARS, Southern Plains Agricultural Research Center, 2765 F&B Road, College Station, TX 77845, USA
JOURNAL

FEATURES
source
1. 279448
Location/Qualifiers
/organism="Sorghum bicolor"
/mol_type="genomic DNA"
/cultivar="I83620C"
/db_xref="taxon:4558"
/chromosome="8"
/map="104.2-111.2 cM"
/clone="BAC 88M4"

ORIGIN

Query Match 4.4%; Score 133.2; DB 8; Length 279448;
Best Local Similarity 77.1%; Pred. No. 2.2e-23;
Matches 162; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Qy 1457 CTAAGGCCAGTCTCAGTGGGTTTTCATCAGATTTTCATGACATTAATAAGCTGATG 1516
|||||
Db 131606 CTAGGTCAGTCTCAATGATGTTTTCATGATTTTAAGGCGATTAAATGCTGATGTA 131665
|||||

gene	similar to rp1"	complement (<27459. .>28260)	/gene="2A"	/evidence=not experimental	complement (join(<27459. .27638,27696. .27866,28135. .>28260))	/gene="2A"	/product="hypothetical protein"	complement (join(27459. .27638,27696. .27866,28135. .28260))	/gene="2A"	/note="similar to EST BE362870"	/codon_start=1	/evidence=not experimental	/product="hypothetical protein"	/protein_id="AAO16707.1"	/db_xref="GI:27542774"	K1IAHRPOLNIMAMTISLQDSQATRSRPSHGLHLPKRDRODSIPATVPGGESFG	CEGGRQTPARGSGSGGAVATSPAAEGGDTLVGLVGCSCVSEHLLTLVES"	complement (<28943. .>29321)	/gene="2B"	/evidence=not experimental	complement (join(<28943. .29038,29166. .>29321))	/gene="2B"	/product="hypothetical protein"	complement (join(28943. .29038,29166. .29321))	/gene="2B"	/codon_start=1	/evidence=not experimental	/product="hypothetical protein"	/protein_id="AAO16708.1"	/db_xref="GI:27542775"	NRQPAKNKHGSDTDGAAGLLLAGKRRKKQGNWKEVE"	<29889. .>30653	/gene="rhp1-4"	/note="rust resistance gene"	/evidence=not experimental	<29889. .>30653	/gene="rhp1-4"	/product="truncated rp1-4 protein"	29889.30653	/gene="rhp1-4"	/note="truncated by retrotransposon insertion; similar to rp1"	/codon_start=1	/evidence=not experimental	/product="truncated rp1-4 protein"	/protein_id="AAO16687.1"	/db_xref="GI:27542754"	/translation="MADIALAGLRWAASPIFNKLINASAYLSVDMVRELQKLEATVLPQDLVIOAAEKSPHKKLEAWRLRLKEAFYDADLLDEHYNLLKRAKSGKDLPLG	EDTSSIIASTMTKPPHTKASRNLLPENRLLISKMNEIKAILTEAKELRDLISLAPG	NTTGLGNVPATIVPTVTSLSKTVKFGDKDRDRIVPLFLKTAADNASSTRYSS	LALIGAGGNGKTSIVQVYNDKRIEEGDIRMVCGV"	30641.33351	/note="contains ITRs"	/transposon="Kaema-1 retrotransposon"	36834.37160	/note="rphi-5; truncated rust resistance gene; similar to rp1"	complement (<40727. .>42969)	/gene="3"	/evidence=not experimental	complement (join(<40727. .42601,42717. .42816,42947. .>42969))	/gene="3"	/product="RPT2-like protein"	complement (join(40727. .42601,42717. .42816,42947. .42969))	/gene="3"	/note="similar to Arabidopsis thaliana RPT2 gene"	/codon_start=1	4.11%; Score 122.8; DB 8; Length 269433; 74.8%; Pred. No. 1.2e-20;
gene	similar to rp1"	complement (<27459. .>28260)	/gene="2A"	/evidence=not experimental	complement (join(<27459. .27638,27696. .27866,28135. .>28260))	/gene="2A"	/product="hypothetical protein"	complement (join(27459. .27638,27696. .27866,28135. .28260))	/gene="2A"	/note="similar to EST BE362870"	/codon_start=1	/evidence=not experimental	/product="hypothetical protein"	/protein_id="AAO16707.1"	/db_xref="GI:27542774"	K1IAHRPOLNIMAMTISLQDSQATRSRPSHGLHLPKRDRODSIPATVPGGESFG	CEGGRQTPARGSGSGGAVATSPAAEGGDTLVGLVGCSCVSEHLLTLVES"	complement (<28943. .>29321)	/gene="2B"	/evidence=not experimental	complement (join(<28943. .29038,29166. .>29321))	/gene="2B"	/product="hypothetical protein"	complement (join(28943. .29038,29166. .29321))	/gene="2B"	/codon_start=1	/evidence=not experimental	/product="hypothetical protein"	/protein_id="AAO16708.1"	/db_xref="GI:27542775"	NRQPAKNKHGSDTDGAAGLLLAGKRRKKQGNWKEVE"	<29889. .>30653	/gene="rhp1-4"	/note="rust resistance gene"	/evidence=not experimental	<29889. .>30653	/gene="rhp1-4"	/product="truncated rp1-4 protein"	29889.30653	/gene="rhp1-4"	/note="truncated by retrotransposon insertion; similar to rp1"	/codon_start=1	/evidence=not experimental	/product="truncated rp1-4 protein"	/protein_id="AAO16687.1"	/db_xref="GI:27542754"	/translation="MADIALAGLRWAASPIFNKLINASAYLSVDMVRELQKLEATVLPQDLVIOAAEKSPHKKLEAWRLRLKEAFYDADLLDEHYNLLKRAKSGKDLPLG	EDTSSIIASTMTKPPHTKASRNLLPENRLLISKMNEIKAILTEAKELRDLISLAPG	NTTGLGNVPATIVPTVTSLSKTVKFGDKDRDRIVPLFLKTAADNASSTRYSS	LALIGAGGNGKTSIVQVYNDKRIEEGDIRMVCGV"	30641.33351	/note="contains ITRs"	/transposon="Kaema-1 retrotransposon"	36834.37160	/note="rphi-5; truncated rust resistance gene; similar to rp1"	complement (<40727. .>42969)	/gene="3"	/evidence=not experimental	complement (join(<40727. .42601,42717. .42816,42947. .>42969))	/gene="3"	/product="RPT2-like protein"	complement (join(40727. .42601,42717. .42816,42947. .42969))	/gene="3"	/note="similar to Arabidopsis thaliana RPT2 gene"	/codon_start=1	4.11%; Score 122.8; DB 8; Length 269433; 74.8%; Pred. No. 1.2e-20;

Matches	154;	Conservative	0;	Mismatches	52;	Indels	0;	Gaps	0;
Qy	1453	GTTTCTAAGGCCAGTCTCAGTGGGGTTTTTCATCAGAGTTCATCGACATTAAATAAGCTGA	1512						
Db	238542	GCTTCTTAGCCCACTCTCAATGAGGGTTTCATTAGAGTTCATGACATTAAATATGCTGA	238483						
Qy	1513	TGTGACACCCTATTTGATGAAGAGAGAGATGATAGAGTTCATGCGCAGTAGAGAGATTT	1572						
Db	238482	TGTGGCACTATATTAAATGAAGAGAGAGATGCTTAAGAGTTCATGGGAGTAGAGAGATTT	238423						
Qy	1573	CATGGGATGAACCTCTTCTTCACTGTTTCCAAAATATAGATGCATTGGTAAGAGGGCCA	1632						
Db	238422	CATCCCATAAACTCTCTGACACTGTTTCCAAAATATTGATGCTTAGAAAACTGTGTCA	238363						
Qy	1633	TGAAATCTCTAGTGACACTGACCTAA	1658						
Db	238362	TGAAACTTCCATTAAGGATGGCCTTA	238337						
RESULT 3									
SCFGLUTRAA									
LOCUS									
DEFINITION	ScFGLUTRAA	1215 bp	mRNA	linear	PLN 02-APR-1999				
SOURCE	Saccharum hybrid cultivar H65-7052 glucose transporter mRNA,								
VERSION	L21752								
KEYWORDS	L21752.1 GI:347852								
ORGANISM	Saccharum hybrid cultivar H65-7052								
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;								
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD								
	clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum								
	complex.								
REFERENCE	1 (bases 1 to 1215)								
AUTHORS	Bugos,R.C. and Thom,M.								
TITLE	Glucose transporter cDNAs from sugarcane								
JOURNAL	Plant Physiol. 103 (4), 1469-1470 (1993)								
MEDLINE	94120022								
PUBMED	8290645								
REFERENCE	2 (bases 1 to 1215)								
AUTHORS	Bugos,R.C.								
TITLE	Direct Submission								
JOURNAL	Submitted (14-JUL-1993) Robert Bugos, Hawaiian Sugar Planters' Association, Experiment Station, Aiea, HI 96701-1057, USA								
FEATURES	Location/Qualifiers								
source	1..1215								
	/organism="Saccharum hybrid cultivar H65-7052"								
	/mol_type="mRNA"								
	/cultivar="H65-7052"								
	/db_xref="taxon:50503"								
	/tissue_type="leaf"								
	<1..866								
	/note="putative"								
	/codon_start=3								
	/product="glucose transporter"								
	/protein_id="AAJ18533.1"								
	/db_xref="GI:347853"								
	/translation="IERGRVEEGRVLERINGTADVDAEFTDMVEASELANITIEHPFR NILEPRNPOLWAVCHPAQLIGINSILFYAPVLFSQMGFGGNASLYSSVLTGAVL FSSTLIISGTVDRGRKKLISGGIOMIVCAIVVAILGAKFGKDKLSRSYSIAVVV VICFLVAFGWSGLGWTPSEIFPLETRSGOSITVANLLFTFAQAFLSLICA PKGFIPLFEAGTIWTMTVCVCFLPETKGVPLEEWVLLWRKHWFVKYMPADMLEDG WGAA PASNNHK"								
CDS	867..1215								
	/note="putative"								
	1215								
	/note="putative"								
3'UTR									
polya_site									

Query Match	4.1%	Score 122.2;	DB 8;	Length 1215;
Best Local Similarity	82.4%	Pred. No. 1.1e-20;		
Matches 164; Conservative	0;	Mismatches 33;	Indels 2;	Gaps 2;

```
QY 1462 GCCAGTCTCAGTGGGGTTTCATCAGAGTTTCATGACATTAATAAGTGTGACACC 1521
      |||
Db 968 GCCAGTCTCAGTGGGGTTTCATCAGAGTTTCATGACATTAATAAGTGTGACACC 1027
      |||
QY 1522 GTATTGATGAGAGAGAGATGATAGAGTTTCATCGAGTAGAGAGTTTCATGG-GGA 1580
      |||
Db 1028 GTATTGATGAGAGAGAGATGATAGAGTTTCATCGAGTAGAGAGTTTCATGGAGGA 1087
      |||
QY 1581 TGAAGTCTCTTCTCACTGTTTCCAAATATAGATGCAATTGGTAAAGAGGGCCCATGAATCT 1640
      |||
Db 1088 TGAAGTCTCTTCTCACTGTTTCCAAATATAGATGCAATTGGTAAAGAGGGCCCATGAATCT 1146
      |||
QY 1641 CTAGTGACACTGACCTAAG 1659
      |||
Db 1147 ATTGAGACTGGTATGTAAG 1165
      |||

RESULT 4
AR412080/c AR412080 3688 bp DNA linear PAT 18-DEC-2003
LOCUS
DEFINITION Sequence 10 from patent US 6638766.
ACCESSION AR412080
VERSION AR412080.1 GI:40164633
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3688)
AUTHORS Albert,H.H. and Wei,H.
TITLE Promoter of the sugarcane UB14 gene
JOURNAL Patent: US 6638766-A 10 28-OCT-2003;
FEATURES
    source 1..3688
    /organism="unknown"
    /mol_type="genomic DNA"

ORIGIN
Query Match 3.9%; Score 118.6; DB 6; Length 3688;
Best Local Similarity 86.0%; Pred. No. 1.1e-19;
Matches 166; Conservative 0; Mismatches 24; Indels 3; Gaps 3;

QY 1451 AGGTTTCTAAGGCCAGTCTCAGTGGGGTTTCATCAGAGTTTCATGACATTAATAAGCT 1510
      |||
Db 1914 AGTTTCTGAGGCCGCTCTCAGT-GGATTTTCATCAGAGTTTCATGACATTAATAAGCT 1856
      |||
QY 1511 GATGTGACACCGTATTGATGAAGAGAGAGATGATAGAGTTTCATCGAGTAGAGAGCT 1570
      |||
Db 1855 GATGTGACACCGTATTGATGAAGAGAGAGATGATAGAGTTTCATCGAGTAGAGAGCT 1796
      |||
QY 1571 TTTCATGGGATGAACTCTT-CTTCACTGTTTCCAAATAT-AGATGCATTGGTAAAGG 1628
      |||
Db 1795 TTTCATGGGATGAACTCTTCTCTGCACTGTTTCCAAATATGGGTTGCATTAATAACATG 1736
      |||
QY 1629 GCCATGAATCTC 1641
      |||
Db 1735 GCCATAAAATCCC 1723
      |||

RESULT 5
AR473380/c AR473380 3688 bp DNA linear PAT 20-FEB-2004
LOCUS
DEFINITION Sequence 10 from patent US 6686513.
ACCESSION AR473380
VERSION AR473380.1 GI:42708831
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3688)
AUTHORS Albert,H.H. and Wei,H.
TITLE Sugarcane ubi9 gene promoter sequence and methods of use thereof
JOURNAL Patent: US 6686513-A 10 03-FEB-2004;
FEATURES
    source 1..3688
    /organism="unknown"
    /mol_type="genomic DNA"

ORIGIN
Query Match 3.9%; Score 118.6; DB 6; Length 3688;
Best Local Similarity 86.0%; Pred. No. 1.1e-19;
Matches 166; Conservative 0; Mismatches 24; Indels 3; Gaps 3;

QY 1451 AGGTTTCTAAGGCCAGTCTCAGTGGGGTTTCATCAGAGTTTCATGACATTAATAAGCT 1510
      |||
Db 1914 AGTTTCTGAGGCCGCTCTCAGT-GGATTTTCATCAGAGTTTCATGACATTAATAAGCT 1856
      |||
QY 1511 GATGTGACACCGTATTGATGAAGAGAGAGATGATAGAGTTTCATCGAGTAGAGAGCT 1570
      |||
Db 1855 GATGTGACACCGTATTGATGAAGAGAGAGATGATAGAGTTTCATCGAGTAGAGAGCT 1796
      |||
QY 1571 TTTCATGGGATGAACTCTT-CTTCACTGTTTCCAAATAT-AGATGCATTGGTAAAGG 1628
      |||
Db 1795 TTTCATGGGATGAACTCTTCTCTGCACTGTTTCCAAATATGGGTTGCATTAATAACATG 1736
      |||
QY 1629 GCCATGAATCTC 1641
      |||
Db 1735 GCCATAAAATCCC 1723
      |||

RESULT 6
AR488131/c AR488131 3688 bp DNA linear PAT 15-MAY-2004
LOCUS
DEFINITION Sequence 10 from patent US 6706948.
ACCESSION AR488131
VERSION AR488131.1 GI:47253896
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3688)
AUTHORS Albert,H.H. and Wei,H.
TITLE Sugarcane UB19 gene promoter and methods of use thereof
JOURNAL Patent: US 6706948-A 10 16-MAR-2004;
FEATURES
    source 1..3688
    /organism="unknown"
    /mol_type="genomic DNA"

ORIGIN
Query Match 3.9%; Score 118.6; DB 6; Length 3688;
Best Local Similarity 86.0%; Pred. No. 1.1e-19;
Matches 166; Conservative 0; Mismatches 24; Indels 3; Gaps 3;

QY 1451 AGGTTTCTAAGGCCAGTCTCAGTGGGGTTTCATCAGAGTTTCATGACATTAATAAGCT 1510
      |||
Db 1914 AGTTTCTGAGGCCGCTCTCAGT-GGATTTTCATCAGAGTTTCATGACATTAATAAGCT 1856
      |||
QY 1511 GATGTGACACCGTATTGATGAAGAGAGAGATGATAGAGTTTCATCGAGTAGAGAGCT 1570
      |||
Db 1855 GATGTGACACCGTATTGATGAAGAGAGAGATGATAGAGTTTCATCGAGTAGAGAGCT 1796
      |||
QY 1571 TTTCATGGGATGAACTCTT-CTTCACTGTTTCCAAATAT-AGATGCATTGGTAAAGG 1628
      |||
Db 1795 TTTCATGGGATGAACTCTTCTCTGCACTGTTTCCAAATATGGGTTGCATTAATAACATG 1736
      |||
QY 1629 GCCATGAATCTC 1641
      |||
Db 1735 GCCATAAAATCCC 1723
      |||

RESULT 7
AR412075/c AR412075 3691 bp DNA linear PAT 18-DEC-2003
LOCUS
DEFINITION Sequence 3 from patent US 6638766.
ACCESSION AR412075
VERSION AR412075.1 GI:40164628
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3688)
AUTHORS Albert,H.H. and Wei,H.
TITLE Sugarcane ubi9 gene promoter sequence and methods of use thereof
JOURNAL Patent: US 6638766-A 10 28-OCT-2003;
FEATURES
    source 1..3688
    /organism="unknown"
    /mol_type="genomic DNA"

ORIGIN
Query Match 3.9%; Score 118.6; DB 6; Length 3688;
Best Local Similarity 86.0%; Pred. No. 1.1e-19;
Matches 166; Conservative 0; Mismatches 24; Indels 3; Gaps 3;

QY 1451 AGGTTTCTAAGGCCAGTCTCAGTGGGGTTTCATCAGAGTTTCATGACATTAATAAGCT 1510
      |||
Db 1914 AGTTTCTGAGGCCGCTCTCAGT-GGATTTTCATCAGAGTTTCATGACATTAATAAGCT 1856
      |||
QY 1511 GATGTGACACCGTATTGATGAAGAGAGAGATGATAGAGTTTCATCGAGTAGAGAGCT 1570
      |||
Db 1855 GATGTGACACCGTATTGATGAAGAGAGAGATGATAGAGTTTCATCGAGTAGAGAGCT 1796
      |||
QY 1571 TTTCATGGGATGAACTCTT-CTTCACTGTTTCCAAATAT-AGATGCATTGGTAAAGG 1628
      |||
Db 1795 TTTCATGGGATGAACTCTTCTCTGCACTGTTTCCAAATATGGGTTGCATTAATAACATG 1736
      |||
QY 1629 GCCATGAATCTC 1641
      |||
Db 1735 GCCATAAAATCCC 1723
      |||
```

```
source 1..3688
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 3.9%; Score 118.6; DB 6; Length 3688;
Best Local Similarity 86.0%; Pred. No. 1.1e-19;
Matches 166; Conservative 0; Mismatches 24; Indels 3; Gaps 3;

QY 1451 AGGTTTCTAAGGCCAGTCTCAGTGGGGTTTCATCAGAGTTTCATGACATTAATAAGCT 1510
      |||
Db 1914 AGTTTCTGAGGCCGCTCTCAGT-GGATTTTCATCAGAGTTTCATGACATTAATAAGCT 1856
      |||
QY 1511 GATGTGACACCGTATTGATGAAGAGAGAGATGATAGAGTTTCATCGAGTAGAGAGCT 1570
      |||
Db 1855 GATGTGACACCGTATTGATGAAGAGAGAGATGATAGAGTTTCATCGAGTAGAGAGCT 1796
      |||
QY 1571 TTTCATGGGATGAACTCTT-CTTCACTGTTTCCAAATAT-AGATGCATTGGTAAAGG 1628
      |||
Db 1795 TTTCATGGGATGAACTCTTCTCTGCACTGTTTCCAAATATGGGTTGCATTAATAACATG 1736
      |||
QY 1629 GCCATGAATCTC 1641
      |||
Db 1735 GCCATAAAATCCC 1723
      |||

RESULT 6
AR488131/c AR488131 3688 bp DNA linear PAT 15-MAY-2004
LOCUS
DEFINITION Sequence 10 from patent US 6706948.
ACCESSION AR488131
VERSION AR488131.1 GI:47253896
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3688)
AUTHORS Albert,H.H. and Wei,H.
TITLE Sugarcane UB19 gene promoter and methods of use thereof
JOURNAL Patent: US 6706948-A 10 16-MAR-2004;
FEATURES
    source 1..3688
    /organism="unknown"
    /mol_type="genomic DNA"

ORIGIN
Query Match 3.9%; Score 118.6; DB 6; Length 3688;
Best Local Similarity 86.0%; Pred. No. 1.1e-19;
Matches 166; Conservative 0; Mismatches 24; Indels 3; Gaps 3;

QY 1451 AGGTTTCTAAGGCCAGTCTCAGTGGGGTTTCATCAGAGTTTCATGACATTAATAAGCT 1510
      |||
Db 1914 AGTTTCTGAGGCCGCTCTCAGT-GGATTTTCATCAGAGTTTCATGACATTAATAAGCT 1856
      |||
QY 1511 GATGTGACACCGTATTGATGAAGAGAGAGATGATAGAGTTTCATCGAGTAGAGAGCT 1570
      |||
Db 1855 GATGTGACACCGTATTGATGAAGAGAGAGATGATAGAGTTTCATCGAGTAGAGAGCT 1796
      |||
QY 1571 TTTCATGGGATGAACTCTT-CTTCACTGTTTCCAAATAT-AGATGCATTGGTAAAGG 1628
      |||
Db 1795 TTTCATGGGATGAACTCTTCTCTGCACTGTTTCCAAATATGGGTTGCATTAATAACATG 1736
      |||
QY 1629 GCCATGAATCTC 1641
      |||
Db 1735 GCCATAAAATCCC 1723
      |||

RESULT 7
AR412075/c AR412075 3691 bp DNA linear PAT 18-DEC-2003
LOCUS
DEFINITION Sequence 3 from patent US 6638766.
ACCESSION AR412075
VERSION AR412075.1 GI:40164628
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3688)
AUTHORS Albert,H.H. and Wei,H.
TITLE Sugarcane ubi9 gene promoter sequence and methods of use thereof
JOURNAL Patent: US 6638766-A 10 28-OCT-2003;
FEATURES
    source 1..3688
    /organism="unknown"
    /mol_type="genomic DNA"

ORIGIN
Query Match 3.9%; Score 118.6; DB 6; Length 3688;
Best Local Similarity 86.0%; Pred. No. 1.1e-19;
Matches 166; Conservative 0; Mismatches 24; Indels 3; Gaps 3;

QY 1451 AGGTTTCTAAGGCCAGTCTCAGTGGGGTTTCATCAGAGTTTCATGACATTAATAAGCT 1510
      |||
Db 1914 AGTTTCTGAGGCCGCTCTCAGT-GGATTTTCATCAGAGTTTCATGACATTAATAAGCT 1856
      |||
QY 1511 GATGTGACACCGTATTGATGAAGAGAGAGATGATAGAGTTTCATCGAGTAGAGAGCT 1570
      |||
Db 1855 GATGTGACACCGTATTGATGAAGAGAGAGATGATAGAGTTTCATCGAGTAGAGAGCT 1796
      |||
QY 1571 TTTCATGGGATGAACTCTT-CTTCACTGTTTCCAAATAT-AGATGCATTGGTAAAGG 1628
      |||
Db 1795 TTTCATGGGATGAACTCTTCTCTGCACTGTTTCCAAATATGGGTTGCATTAATAACATG 1736
      |||
QY 1629 GCCATGAATCTC 1641
      |||
Db 1735 GCCATAAAATCCC 1723
      |||
```

```
KEYWORDS      Unknown.
SOURCE        Unknown.
ORGANISM      Unclassified.
REFERENCE     1 (bases 1 to 3691)
AUTHORS      Albert,H.H. and Wei,H.
TITLE        Promoter of the sugarcane UB14 gene
JOURNAL      Patent: US 6638766-A 3 28-OCT-2003;
FEATURES     Location/Qualifiers
            source
            1..3691
            /organism="unknown"
            /mol_type="genomic DNA"

ORIGIN
Query Match      3.9%; Score 118.6; DB 6; Length 3691;
Best Local Similarity 86.0%; Pred. No. 1.1e-19;
Matches 166; Conservative 0; Mismatches 24; Indels 3; Gaps 3;

QY 1451 AGGTTTCTAAGCCAGTCTCAGTGGGGTTTCATCAGAGTTTCATGCACATTAAATAAGCT 1510
    |||||
Db 1914 AGTTTCTGAGGCCGCTCTCAGT-GGATTTTCATCAGAGTTTCATGCACATTAAATAGGCT 1856

QY 1511 GATGTGACACCGTATTGATGAAGAGAGAGATGATAAGAGTTTCATGCCAGTAGAGAGT 1570
    |||||
Db 1855 GATGTGGCACCGTATTGATGAAGAGAGAGATGATAAGAGTTTCATGGAATTAGAGAGT 1796

QY 1571 TTCATGGGATGAAACTCTT-CTTCACTGTTTCCAAAATAT-AGATGCATTGGTAAGAGG 1628
    |||||
Db 1795 TTTACGAAGATGAAACTCTTCTCGACTGTTTCCAAAATATGGGTTCATTAAATACATG 1736

QY 1629 GCCATGAATCTC 1641
    |||||
Db 1735 GCCATAAAATCCC 1723

RESULT 8
AR473375/c    AR473375    3691 bp    DNA    linear    PAT 20-FEB-2004
LOCUS         AR473375
DEFINITION    Sequence 3 from patent US 6686513.
ACCESSION     AR473375
VERSION       AR473375.1 GI:42708826
KEYWORDS      Location/Qualifiers
SOURCE        1..3691
              /organism="unknown"
              /mol_type="genomic DNA"

ORGANISM      Unclassified.
REFERENCE     1 (bases 1 to 3691)
AUTHORS      Albert,H.H. and Wei,H.
TITLE        Sugarcane ubi9 gene promoter sequence and methods of use thereof
JOURNAL      Patent: US 6686513-A 3 03-FEB-2004;
FEATURES     Location/Qualifiers
            source
            1..3691
            /organism="unknown"
            /mol_type="genomic DNA"

ORIGIN
Query Match      3.9%; Score 118.6; DB 6; Length 3691;
Best Local Similarity 86.0%; Pred. No. 1.1e-19;
Matches 166; Conservative 0; Mismatches 24; Indels 3; Gaps 3;

QY 1451 AGGTTTCTAAGCCAGTCTCAGTGGGGTTTCATCAGAGTTTCATGCACATTAAATAAGCT 1510
    |||||
Db 1914 AGTTTCTGAGGCCGCTCTCAGT-GGATTTTCATCAGAGTTTCATGCACATTAAATAGGCT 1856

QY 1511 GATGTGACACCGTATTGATGAAGAGAGAGATGATAAGAGTTTCATGCCAGTAGAGAGT 1570
    |||||
Db 1855 GATGTGGCACCGTATTGATGAAGAGAGAGATGATAAGAGTTTCATGGAATTAGAGAGT 1796

QY 1571 TTCATGGGATGAAACTCTT-CTTCACTGTTTCCAAAATAT-AGATGCATTGGTAAGAGG 1628
    |||||
Db 1795 TTTACGAAGATGAAACTCTTCTCGACTGTTTCCAAAATATGGGTTCATTAAATACATG 1736

QY 1629 GCCATGAATCTC 1641
    |||||
Db 1735 GCCATAAAATCCC 1723

RESULT 9
AR488126/c    AR488126    3691 bp    DNA    linear    PAT 15-MAY-2004
LOCUS         AR488126
DEFINITION    Sequence 3 from patent US 6706948.
ACCESSION     AR488126
VERSION       AR488126.1 GI:47253891
KEYWORDS      Location/Qualifiers
SOURCE        1..3691
              /organism="unknown"
              /mol_type="genomic DNA"

ORGANISM      Unclassified.
REFERENCE     1 (bases 1 to 3691)
AUTHORS      Albert,H.H. and Wei,H.
TITLE        Sugarcane UB19 gene promoter and methods of use thereof
JOURNAL      Patent: US 6706948-A 3 16-MAR-2004;
FEATURES     Location/Qualifiers
            source
            1..3691
            /organism="unknown"
            /mol_type="genomic DNA"

ORIGIN
Query Match      3.9%; Score 118.6; DB 6; Length 3691;
Best Local Similarity 86.0%; Pred. No. 1.1e-19;
Matches 166; Conservative 0; Mismatches 24; Indels 3; Gaps 3;

QY 1451 AGGTTTCTAAGCCAGTCTCAGTGGGGTTTCATCAGAGTTTCATGCACATTAAATAAGCT 1510
    |||||
Db 1914 AGTTTCTGAGGCCGCTCTCAGT-GGATTTTCATCAGAGTTTCATGCACATTAAATAGGCT 1856

QY 1511 GATGTGACACCGTATTGATGAAGAGAGAGATGATAAGAGTTTCATGCCAGTAGAGAGT 1570
    |||||
Db 1855 GATGTGGCACCGTATTGATGAAGAGAGAGATGATAAGAGTTTCATGGAATTAGAGAGT 1796

QY 1571 TTCATGGGATGAAACTCTT-CTTCACTGTTTCCAAAATAT-AGATGCATTGGTAAGAGG 1628
    |||||
Db 1795 TTTACGAAGATGAAACTCTTCTCGACTGTTTCCAAAATATGGGTTCATTAAATACATG 1736

QY 1629 GCCATGAATCTC 1641
    |||||
Db 1735 GCCATAAAATCCC 1723

RESULT 10
AR412079/c    AR412079    5174 bp    DNA    linear    PAT 18-DEC-2003
LOCUS         AR412079
DEFINITION    Sequence 8 from patent US 6638766.
ACCESSION     AR412079
VERSION       AR412079.1 GI:40164632
KEYWORDS      Location/Qualifiers
SOURCE        1..5174
              /organism="unknown"
              /mol_type="genomic DNA"

ORGANISM      Unclassified.
REFERENCE     1 (bases 1 to 5174)
AUTHORS      Albert,H.H. and Wei,H.
TITLE        Promoter of the sugarcane UB14 gene
JOURNAL      Patent: US 6638766-A 8 28-OCT-2003;
FEATURES     Location/Qualifiers
            source
            1..5174
            /organism="unknown"
            /mol_type="genomic DNA"

ORIGIN
Query Match      3.9%; Score 118.6; DB 6; Length 5174;
Best Local Similarity 86.0%; Pred. No. 1.1e-19;
Matches 166; Conservative 0; Mismatches 24; Indels 3; Gaps 3;

QY 1451 AGGTTTCTAAGCCAGTCTCAGTGGGGTTTCATCAGAGTTTCATGCACATTAAATAAGCT 1510
    |||||
Db 1914 AGTTTCTGAGGCCGCTCTCAGT-GGATTTTCATCAGAGTTTCATGCACATTAAATAGGCT 1856

QY 1511 GATGTGACACCGTATTGATGAAGAGAGAGATGATAAGAGTTTCATGCCAGTAGAGAGT 1570
    |||||
Db 1855 GATGTGGCACCGTATTGATGAAGAGAGAGATGATAAGAGTTTCATGGAATTAGAGAGT 1796

QY 1571 TTCATGGGATGAAACTCTT-CTTCACTGTTTCCAAAATAT-AGATGCATTGGTAAGAGG 1628
    |||||
Db 1795 TTTACGAAGATGAAACTCTTCTCGACTGTTTCCAAAATATGGGTTCATTAAATACATG 1736

QY 1629 GCCATGAATCTC 1641
    |||||
Db 1735 GCCATAAAATCCC 1723
```



```

/translation="MQIFVKTLLTGKTTITLVESSDTIDNVKAKIQDKGIPDQORLI
PAGQLEDGRTLDYNIQKSTLHLVLRGGMQIFVKTLLTGKTTITLVESSDTIDNV
KAKIQDKGIPDQORXIFAGQLEDGRTLDYNIQKSTLHLVLRGGMQIFVKTLL
TGKTTITLVESSDTIDNVKAKIQDKGIPDQORLIIPAGQLEDGRTLDYNIQKST
LHLVLRGGMQIFVKTLLTGKTTITLVESSDTIDNVKAKIQDKGIPDQORXIFAG
QLEDGRTLDYNIQKSTLHLVLRGGMQIFVKTLLTGKTTITLVESSDTIDNVKAKI
QDKGIPDQORLIIPAGQLEDGRTLDYNIQKSTLHLVLRGGMQIFVKTLLTGKTT
4835..5079
/gene="ubi9"

ORIGIN
Query Match      3.9%; Score 118.6; DB 8; Length 5174;
Best Local Similarity 86.0%; Pred. No. 1.1e-19;
Matches 166; Conservative 0; Mismatches 24; Indels 3; Gaps 3;

QY 1451 AGTTTCTAAGCCAGCTCTCAGTGGGTTTCATCAGAGTTTCATGACATTAAATAGCT 1510
Db 1914 AGTTTTCTGAGGCCCGTCTCAGT-GGATTTTCATCAGAGTTTCATGACATTAAATAGCT 1856

QY 1511 GATGTGACACCGTATTGATGAAGAGAGAGATGATAAGAGTTTCATGCCAGTAGAGAGAGT 1570
Db 1855 GATGTGGCACCGTATTGATGAAGAGAGAGATGATAAGAGTTTCATGGAATTAGAGAGAGT 1796

QY 1571 TTCATGGGAGTAAACTCTT-CTTCACCTGTTTCCAAATAT-AGATGCATTGGTAAGAGG 1628
Db 1795 TTACGAAGATGAACACTCTTCTGCTGACTGTTTCCAAATATGGGTGCATTAAATACATG 1736

QY 1629 GCCATGAATCTC 1641
Db 1735 GCCATAAATCCC 1723

RESULT 14
AF114171 183990 bp DNA linear PLN 25-APR-1999
LOCUS Sorghum bicolor BAC clone 25.M18, complete sequence.
DEFINITION AF114171
ACCESSION AF114171
VERSION AF114171.1 GI:4680196
KEYWORDS Sorghum bicolor (sorghum)
SOURCE Sorghum bicolor
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 183990)
AUTHORS Llaça,V., Lou,A., Young,S. and Messing,J.
TITLE Retrotransposable elements of Sorghum bicolor
JOURNAL unpublished
REFERENCE 2 (bases 1 to 183990)
AUTHORS Llaça,V., Lou,A., Young,S. and Messing,J.
TITLE Direct Submision
JOURNAL Submitted (15-DEC-1998) Wakman Institute, Rutgers University, 190
Frelinghuysen Rd., Piscataway, NJ 08854-8020, USA
FEATURES
source
1..183990
/organism="Sorghum bicolor"
/mol_type="genomic DNA"
/db_xref="taxon:4558"
/chromosome="P"
/clone="BAC 25.M18"
/location="Woo et. al., Nucl. Acids Res. 22:4922-4931,"
/repeat_region 3517..3644
/rpt_family="HCSR-1"
/rpt_type="dispersed"
join(<3841..4596,4699..5541)
/product="putative beta-ketoacyl-CoA synthase"
join(3841..4596,4699..5541)
/note="similar to Oryza sativa EST clones E10604_1A,
E31307_5A, S3416_3A"
/codon_start=1
/product="putative beta-ketoacyl-CoA synthase"
/protein_id="AAD27560.1"
/db_xref="GI:4680197"

```

```

/translation="MSIKSETNRSSTIQSAGHRICSTVLCYQHOHLMAVSYALVFL
APILATLHLSHRTSADDLLIYSARSTLTANLPFTVVVAAATAYLLMRPRAVY
LLDIACKPGNEHVAETENRVLVVGTFESDDSLAFOFKILNRSGLQCTYLPKALT
SSPNPCMAERAEAEAVFAGDQLLAKTQVRAQDGIIVVNSGLFSTPSLAMIY
HMLKRDVSSYNLAGGCCSASLISVDLAKQLLQVQNTYALVWSLEIMAAASYGNT
RSMVLNPCIIFRMGAALVLTNHTERRAKYQLLHTVTHQADDRAYRCVYQEBDES
GHVGVALSOLDMSVAALAKTNTITLGPRLVPMSEKMKMLVGGKAAVVQVLGVQOQG
QGSRSRPYIPDLTAFEHFCIHAGGRAVLDTVERSLDSAWHMEPSSMTLYRWGNTS
SSSLMYELAYTEAKGRVRRGRVCOIAPGSGFKCNSAVWRALRTIDPAKESAAGNPW
VEIHRPVEVPQVERVDVAMAS"
complement(5825..5905)
/rpt_family="HCSR-4"
/rpt_type="dispersed"
complement(7346..8203)
/gene="22-kDa kafirin"
/pseudo
11416..11598
/rpt_family="HCSR-1"
/rpt_type="dispersed"
join(<11990..12670,13053..13158,14484..14704,
14766..>14987)
/product="hypothetical protein"
join(11990..12670,13053..13158,14484..14704,14766..14987)
/codon_start=1
/product="hypothetical protein"
/protein_id="AAD27578.1"
/db_xref="GI:4680215"
/translation="MYTPYDSSSLKTLNGQATSDHACTLNFPQIYQIPESRATT
NVHRPLDRALLAVTAAAGVAMALHRDVLVHRRGELRVAIPIRRSDVAPKHITGEBD
ESKEVLEGLGALLRAGKAVADVAVGVTRAPGGHGTVERERLVEPEPLQRDSDAR
GEVHREARQROAQDGGDAERHEDDEVVVAFETHVRRGVEVEHQRRCPAPAAV
AECVAGROEHSNAKQSOTATATLVLSKSPATHLAPLTKCVGIKEARGRCAPAAAV
PDRRRRTGRCGQTDRPGPELGGEGESAESEEHHPAVLEVVEVRVDFPAVDED
PSSSRSGEEVVVTRFHTARNVDGAAACRAEHQTRHRTSRNSVPGSTWPLND
PMVVRPDGRMDRLVD"
13436..13657
/rpt_family="HCSR-1"
/rpt_type="dispersed"
13994..14148
/rpt_family="HCSR-1"
/rpt_type="dispersed"
complement(15740..15838)
/rpt_family="HCSR-1"
/rpt_type="dispersed"
complement(15910..16127)
/rpt_family="HCSR-1"
/rpt_type="dispersed"
16501..16675
/rpt_family="HCSR-1"
/rpt_type="dispersed"
17050..17249
/rpt_family="HCSR-1"
/rpt_type="dispersed"
complement(17650..17893)
/rpt_family="HCSR-1"
/rpt_type="dispersed"
complement(19738..19888)
/rpt_family="HCSR-4"
/rpt_type="dispersed"
20094..20411
/rpt_family="HCSR-1"
/rpt_type="dispersed"
complement(22849..23088)
/rpt_family="HCSR-1"
/rpt_type="dispersed"
complement(23211..23402)
/rpt_family="HCSR-1"
/rpt_type="dispersed"
complement(join(<26698..27270,27361..27761,27976..28284,
29824..31036,31131..31286,31385..31568,32230..32756,
32998..33453,33728..34017,34138..34292,34882..353045))
/product="hypothetical protein"
complement(join(26698..27270,27361..27761,27976..28284,

```



```
29824..31036,31131..31286,31385..31569,32230..32756,
32998..33453,33728..34017,34138..34292,34882..35045))
/codon_start=1
/product="hypothetical protein"
/protein_id="AAD27577.1"
/db_xref="GI:4680214"
/translation="MDGQAAARAAQGVPGVADLVAAALRQLKFLAAVDRRWLVLD
ECLNLSLVLIQIAHIFQVQYRDCKVGRILNNVSSSTQKSLIQSEKIMKELYP
LPPKLEFKTSDVMDVPGVAEDITVDLSAVKQSSFYQVGTVMHDSRFLQEA
LARYKAFILTKMNEKGLQRPVPTVDVLLMTHOLHPVTVYENDVVKLGLLEHD
D'DADRSSEKLDVGTTFETQFENFGRVTKWAGCMYRGNPSPVSTPQIFSTEVG
TGSDDCKAQNALDITAVELYLOIVDINLPSAVKENYVFRTPNQDPIFSDGG
KLDISTVTKNAGVLCQTEGLILVMDVQSKPEPIGKVFVPMWEPFLSKCLPHSIRK
PKFWELKAGHGATSPPVSLRVAASATVPSFQVFSVMWEPFLSKCLPHSIRK
TLKDSMLSTHGTGDSLLDGDGNOLIKLYOGRKLEYERKCSAHSDEVSATVAKFC
ABHPYKGAVALDTSQILMNEDEWFLPWTTSVLEFMDAGDKDSMLIGAVVQKGA
VFGSDTSVLETFQAVAGSVATAPACQCTCFSTAFSGDKVAVGAFKRAHSSASRTVV
SGENHTESASCGCGGCGGSCGPMVAEDKADNAGCGCGGCGGCGGCGGCGGCGG
ASTMASGQSGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
GCGGCGMAIEGSKTANHAHSGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
RSGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
AGEGKHVSQGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
QTRERADGQGGSLAAKRLIARVILIKARVDGRRRESRDLVRPSLYLSLSPWT
PSSVGTPLFSGNVQAQIEDVESQPDVTGRRKAKVTSRQKQKFDIEEDQVVC
SAWLNVDKPIINGANQGTFWGRIRAFPEKHDKKTAQCVRTRESSIMRWLTITQDV
NKFQSCVKEIKRNASGATIDQMINAALETYICADEENKPTFLAHKYLEEKEDKQ
LRLEAKNKKNTKTESPTNVNEASNDIEIPEVADPSEERKPKGQYKQVQKARRDD
ASLAVLEKMLEKEARELEREDKAREASLEIEASLEIEKALEIEKRLSNEEKIEA
ELWKEKDIIMLADKSLDDEQLQWLQIMKNKILARHTGN"
36421..36587
/repeat_region
/rpt_family="HCSR-4"
/rpt_type="dispersed"
/complement(join(38130..38275,38407..38612,38705..38895,
39620..39776,39932..40049,40120..40673,40764..40864,
41122..41697,41780..42115,42344..42727,42847..43083))
/product="hypothetical protein"
/complement(join(38130..38275,38407..38612,38705..38895,
39620..39776,39932..40049,40120..40673,40764..40864,
41122..41697,41780..42115,42344..42727,42847..43083))
/codon_start=1
/product="hypothetical protein"
/protein_id="AAD27562.1"
/db_xref="GI:4680199"
/translation="MKTSLARVASVEFLPLVLGFLNLLQLQSKANQVRQKLOQL
LPKSNFYCTISKAPPCFHLRLNDQVGENYPLPGSSAATIGSTAASPTAAARR
RPAASQASVATPTRAHGRDWLGLRRHRCRLDDGLRPLTRGRCSKTKAGISRD
RGCTSTRRLHLRLRRGRSVSDAAASLPDRLPKGSRSLGRKIEIFLFGSMG
DRADWDRFLNLPDACKSELDAGNPMGIPTATGWKNVSKFGDKGDKRTKKLN
KLDILKRYTFMFKNFATGLGWAENKTVPCKEWDEHLARCNPEKIKCSHVK
FRKHGKFLDMHVIQKAGVDSGTATCPDTSDDSDSDVAEVPKPKTEKTVNQ
GVRKGTSTAAEKEEKEKSPFFMYNTCLKIENAAERISTSVSSAATINQIFISA
EAVMKVECEBGTALMTAWALANPEFEFTSLETNGLRLDIOREHEKRNKK
EGSMDDHLRAVAGQGLLILAEISQAAIYVGMGNSFYKFMRLSELYENHLLV
SNYGLTSKNVSIESLAEFLMTVGPQSPQAEENFTSLMTIHTKPEHLECLRLK
AKNITPKDPTFSTEHEHVEENFMPYFKGAIQADISGHVKVITPLENTVNTNRGY
TSQNLAICDFRFTFVATGNGSADSRILNHALANPSPFPVAPKGYLVDSGYA
NRNGYLAFFKGSYTHIDFERNSEPPQGNRRDKEDFRCDLNRGGGSRVQAVRSHGL
DGTDEGSDSMNIIRDIVDALTKETRAARRAEDSLNTRGGGSRVQAVRSHGL
RRPLVVGGRRRRRRAGVQSHVQSPRICHTAHLQRLPLPQGGQIQOAVHVQVRA
AERAGRRRTTRVAGRGHGLRLDRDRRHRQLPPELRRGYSAGADARRVGAR
RRPRLIPAPVRRRLAPARLWNTHRIAM"
join(<33647..43908,43994..44222,44944..45484)
/product="hypothetical protein"
/join(43647..43908,43994..44222,44944..45484)
/codon_start=1
/product="hypothetical protein"
/protein_id="AAD27563.1"
/db_xref="GI:4680200"
/translation="MWSGQIDDSLIWLSVRMIPRYKOVLSFLFGWSHGIRCOHPSID
LVRLDQFSREQTREQGYRTNDLRFVEEQHCRIASDSDVCTAQEAQARSADAVGLHPPR
```

```
LSIDLINISETAFVASLLHPESRAALATAPSSCSLAAASGTSQAASSTLKRWPSW
IDRATSFITFTVYIERLTCMEIHKXDQDQIRSLRPPQDARVAVNFKIDEVYPSR
DNARLSDNIQNTNIESPSPFLVQIAPRPPLPSPFVSCNLCRPGAHVACALCVWRG
RCRCRRNDNTGRQMLFMPADYTHVTVIYDDDSIIHALLLVLLACQQLALC
TLVYHLRCT"
complement(45613..45755)
/rpt_family="HCSR-4"
/rpt_type="dispersed"
/complement(join(46323..46753,47300..47458,48167..48248))
/product="hypothetical protein"
/complement(join(46323..46753,47300..47458,48167..48248))
/notes="similar to Oryza sativa EST clone C3025_1A"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAD27561.1"
Query Match 3.5%; Score 106.4; DB 8; Length 183990;
Best Local Similarity 76.9%; Pred. No. 2.6e-16;
Matches 143; Conservative 0; Mismatches 41; Indels 2; Gaps 1;
mRNA
CDS
repeat_region
1439 GGGGTAGGAGTAGTTTCTTAAGCCAGTCTAGTCGGGTTTCATCAGAGTTTCATGGAC 1498
161654 GTGCCCGCATATCACTGTTAGCCAGGCTTAATAGTTTATCATGTTTCATAGAC 161713
1499 ATTAATAAGCTGATGTGACACCGTATTGATGAAGAGAGAGATGATAAGAGTTTCATGCG 1558
161714 ACTAATAATGCTGATGTGACACTATTAATGAAGAGAGAGGTGATAAGAGTTTCATAAG 161773
1559 AGTAGAGAGAGTTTCATGGGATGAACTTTCTTCACTGTTTCCAAATATAGATGCAT 1618
161774 AGT--AGAGAGTTTATAAGGATGAACTCATCTGCACATATTTCCAAATATGAGGTGT 161831
1619 TGGTAA 1624
Db 161832 TGAANA 161837
RESULT 15
AF114171/c 183990 bp DNA linear PLN 25-APR-1999
LOCUS Sorghum bicolor BAC clone 25.M18, complete sequence.
DEFINITION AF114171
ACCESSION AF114171
VERSION AF114171.1 GI:4680196
KEYWORDS
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
AUTHORS 1 (bases 1 to 183990)
Llaca,V., Lou,A., Young,S. and Messing,J.
TITLE Retrotransposable elements of Sorghum bicolor
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 183990)
Llaca,V., Lou,A., Young,S. and Messing,J.
AUTHORS Direct Submission
TITLE Submitted (15-DEC-1998) Waksman Institute, Rutgers University, 190
JOURNAL Frelinghuysen Rd., Piscataway, NJ 08854-8020, USA
FEATURES
source
1..183990
/organism="Sorghum bicolor"
/mol_type="genomic DNA"
/db_xref="taxon:4558"
/chromosome="F"
/clone="BAC 25.M18"
/clone_lib="Woo et. al., Nucl. Acids Res. 22:4922-4931,"
3517..3644
/rpt_family="HCSR-1"
/rpt_type="dispersed"
/join(<3841..4596,4699..45541)
/product="putative beta-ketacyl-CoA synthase"
/join(3841..4596,4699..45541)
/notes="similar to Oryza sativa EST clones E10604_1A,"
```



```
E31307.5A, S3416.3A"
/codon_start=1
/product="putative beta-ketoacyl-CoA synthase"
/protein_id="AAD27560.1"
/db_xref="GI:4680197"
/translation="MSIKSETNRSSTQTSAGHRCISVGLCYOHLMLAVSYALYLF
APILATLVLSRHTSADDDLIIVSARSTLTANLPFTVVVAAAATVLMRRPRAY
LLDLAKYKPAEHAHVTRTFMRLVVGTFSDSLAFORFKILERSGLGOGYLPAKLT
SSPNPCWAEARRAEAEVMEGAVDQLLAKTGVRAQDGIIVVNSGLSPPTSLSAMIV
HHYKLRDVSSYNLGMCSSASLISVDLAKQLQVQRTYALVLSLEIMASAYGNT
RSMVLVPCIFRMGGAUULLTNHTRERRAKTQLLHVTHQGDADRAVCYQBEDES
GGVVALSKDMLVAAAEKHTNITLGRVLPVSEKMMVLGGKAAVVLVGVQOQG
QGGSRFYIPDLTAFHFCIHAGGRAVLDLVERSLLSAAHMEPSMTLYRWGQS
SSSLYELAYTEAKGRVRRHRVCOIAFGSGFKCNASVWRALRTIDPAKESAAAGNPW
VEETHRPVEVPOVERVDVAAMAS"
complement(5825..5905)
/rpt_family="HCSR-4"
/rpt_type=dispersed
complement(7346..8203)
/gene="22-kDa kafirin"
/pseudo
repeat_region
11416..11598
/rpt_family="HCSR-1"
/rpt_type=dispersed
join(<11990..12670,13053..13158,14484..14704,14766..14987)
/product="hypothetical protein"
/protein_id="AAD27578.1"
/db_xref="GI:4680215"
/translation="WYTPYDPSESILKTLNGATSDHACTLNFPIQLYQIOEPSRATT
NVHRLDRLALLATVAGVANALHRRDVLVVRGELRVAIPRRSDSVAPKHITGEED
ESKLEVEGLLGLALRGAMADVAVGVRAPGGEHGVTERRLRVLPEQRODEAR
GEVHREARQRAQGGDGADEHDEEVVAFETHVRGVGVVEHQLRERRRYV
AECVAGQREHNAQSOTALATVLSSSKPAFLPRTKVGIVEARGRCPPAAAV
TDDRRRRRTGRGQTRPGPELGGEGSEAEHEHPAVLEVVEVRRVDFPAVDED
PESSSRSGREVVVTRTFHARTNVGDGAACRADEHQTTRHPTSRNSVPGSTVPLND
PMVVRPDRGMDRLVDD"
repeat_region
13436..13657
/rpt_family="HCSR-1"
/rpt_type=dispersed
repeat_region
13994..14148
/rpt_family="HCSR-1"
/rpt_type=dispersed
repeat_region
complement(15740..15838)
/rpt_family="HCSR-1"
/rpt_type=dispersed
repeat_region
complement(15910..16127)
/rpt_family="HCSR-1"
/rpt_type=dispersed
repeat_region
16501..16675
/rpt_family="HCSR-1"
/rpt_type=dispersed
repeat_region
17050..17249
/rpt_family="HCSR-1"
/rpt_type=dispersed
repeat_region
complement(17650..17893)
/rpt_family="HCSR-1"
/rpt_type=dispersed
repeat_region
complement(19738..19888)
/rpt_family="HCSR-4"
/rpt_type=dispersed
repeat_region
20094..20411
/rpt_family="HCSR-1"
/rpt_type=dispersed
repeat_region
complement(22849..23088)
/rpt_family="HCSR-1"
/rpt_type=dispersed
repeat_region
complement(23211..23402)
/rpt_family="HCSR-1"
/rpt_type=dispersed
```

mRNA

CDS

repeat_region

mRNA

CDS

mRNA

CDS

```
complement(join(<26698..27270,27361..27761,27976..28284,
29824..31036,31131..31286,31385..31568,32230..32756,
32998..33453,33728..34017,34138..34292,34882..35045))
/product="hypothetical protein"
complement(join(26698..27270,27361..27761,27976..28284,
29824..31036,31131..31286,31385..31568,32230..32756,
32998..33453,33728..34017,34138..34292,34882..35045))
/codon_start=1
/product="hypothetical protein"
/protein_id="AAD27577.1"
/db_xref="GI:4680214"
/translation="MDGEQARAAQAQEGVPVGADLVAAALRQLKFLAARDRRRLYD
EGPLLRATIRRYKACPLDKDKTKAAAVDGLVPLVDCIWHCHRNLVHILTMFI
ENSQLSLITIOAHIFOVYTRDCKVYTRLLNNNVESSTQTSILQSEKIKWELYP
EEPKLFSTKTSDDVMDVNGVAAEDITVDLVAVKROSSFYQVGTTHMDSRLEQA
LARYAKFLYLKMMQEKQLRFRVPTVDLLHQLHPVYTRNDVKKLQKVLJLHED
DTDADRSEKKLDVGTFTTEQFENTGVRYKAGMYRGNMPSVTSPTQIFSTFVG
TGSDICRAQDINALDITAVELYQIVDINNPSAVRKNYVYVFTQNPDTFTSDGG
KLDISTVTKNAGVLCQCEPTGELILVVMVDQVSKPPIKGVKPSLQDLIGPDSKLS
FEKWFELKAGGHATSPVSLRVAASATVPSSFQKVFMSVMMEPSLKCILPHSIKD
QNMSSWTRFYVDCGTELIRLQIRBQAKMGAAVRELGVLSKPKQFQPLAEFKNKW
TLKDSMLSIHTGTDGSLDLKGNQLIKLYQGRKLEYERKCCSHSESDVSAVTKFC
AEHPYKGAVALLDTEFSQLIMVNDWFLPMTTTSVLFMDADGDKMLIGGAAVQKGA
VFGSDTSMVLETOAGAGSVATAQACGTCTAFSGDKVAVGAEHASSGASRTVV
SGENSHTESASGCGGCGGCPNVAEDSKADNAGCGCGCGCGCGCGCGCGTMLK
ASTWASRQSRSGCGGCGGCGGCGGCGTMLKASTWASRQSRSGCGGCGGCGGCG
GCGSGGCGGCGGMAIEGSKTANHAKSGCGGCGGCGGCGGCGTGLFSSSTAADGHS
RSSGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGTGFNVSTK
AGRGHSHSVGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG
QTRRADGQSGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG
PSSSVGTPFGSNVVAQLOEDVVEVSDVPTVGVGRRAKKTACVCTESSINHWLTJORDV
SAWLNSKQDPIGANGQGRSTFMWGRIRAFEFKHKDKKTACVCTESSINHWLTJORDV
KNFCSCEKIERNNASGATQDMINAALETYIGADEENKPTLHCYILKLEEDKXKS
LRLEAKXKXKQKTTNESTPVEASNNDEIPEVAAPDSERKPKQKQKQKQKQKQKQKQK
ASLAVLEKMLEKKEARELERDKAREASLETERASLETERASLETERASLETERASLE
ELMKEEDIMLADKTSLDEQLQWLMKNKILARHTGN"
36421..36587
/rpt_family="HCSR-4"
/rpt_type=dispersed
complement(join(<38130..38275,38407..38612,38705..38895,
39620..39776,39932..40049,40120..40673,40764..40864,
41122..41697,41780..42115,42344..42727,42847..43083))
/product="hypothetical protein"
complement(join(38130..38275,38407..38612,38705..38895,
39620..39776,39932..40049,40120..40673,40764..40864,
41122..41697,41780..42115,42344..42727,42847..43083))
/codon_start=1
/product="hypothetical protein"
/protein_id="AAD27562.1"
/db_xref="GI:4680199"
/translation="MKTSACLARVAVFRLPLVGLFNLLLLQKSKANQRVKQLQQL
LPKSNFYCSTISKAPPFCFRLNDQGENYPLPFGSSAATIGTSTAASPATAAARR
RPAASQAASPVATPTRAHGRDWLGRFLRRHCLRLDDGRLPTRGCGSKAGCISRRD
RGCISTRRLRLRLPDKGGRSYSDAAAASIFDDDLRPGKSRGKRIEILFGSSMG
DKADSDRFLRNLFDACEEIDAGNPMGIPTATGMKNVSKFGDKGDKRTKKOLK
KLDILKKEYTMFMEFKNPAFGLDWDAENKTVDCPEKWDHLEARNPEKIKCSHVX
FRKGPKFLSDMHVIFGKAHVDMGATPCGDISDSDSDSDSDSDSDSDSDSDSDSDSD
EAVRWKCEGVGEETGLTAWALANPEFEFTFSTLTNNRDLIQREHEKKNKK
EGSMDDHLRAVAKQKGLLILAELOSHAAIVGGMGNSYFKMFMLSLVEFNEHDLV
SNYGLTSTKNVSSIESLAFMLWVGGPQSFQANRFTSLWTIHTKFEHVLRLKRL
AKONITKPDPTFTEHERVRENRPYKGAIGADQSHVAVIYPLNETVNTHTRGY
TSQNVLAICDFDMRFTFVATGPGSADSLNLHALANPFPVPAKGYLYLVDGSGYA
NRNGYLAFFKSTYTHIPEFNRSEPPQGNKMRDEFCDDADEDMPCARRRTTHE
DGTDEGESDESMNIRDIRIDALTKEKTRARRAEDSLNTRGGGSRYPQAVRSHGAL
RRPLVVGSGRRRRRGGAVVQSSHVQPICTAHLQOEPRLVQOQIQQVAVHGVRA
AREAGRRRTTRVAGPRGHGRRLRRRRGRHQRLLPELRLRGYSGDAGACAGVGAR
RRPRLLPAPVRRRLARLNMNTHIAM"
join(<43647..43908,43994..44222,44944..45484)
/product="hypothetical protein"
join(43647..43908,43994..44222,44944..45484)
/codon_start=1
```

Wed Feb 16 17:07:28 2005

us-10-751-612-1.rge

```
/product="hypothetical protein"
/protein_id="AAD27561.1"
/db_xref="GI:4680200"
/translation="MSGGIQDDSLIWLVRMIPRYKQVLSFLGWSHGIRCOHPSID
IVRLPQFREGTREGYNTDLFVEEQHCRRIASDVCTAQEAQRASADAVGLHPPR
LSLDLNISETFAVSLHHPESAALATARSCLAAASGSAQASSTLKRWFSW
IDRATSFITTTVTVEIRLIGMEIHCKDVQDQEISSLRPPQDFARVAVNFKIDVYPSR
DNAVLSDPNICTIESPVFLVQIAPRPPPLSPVCSRNLCPGFAHVACALCVWRG
RCRCRRNDNTGRQMLIFMPADYYTHVTVIIYDDDDSIHALLLVLLACGQLLALC
TLVHLRCT"
/complement(45613..45755)
/rpt_family="HCSR-4"
/rpt_type="dispersed"
/product="hypothetical protein"
/complement(join(46323..46753,47300..47458,48167..48248))
/notes="similar to Oryza sativa EST clone C3025_1A"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAD27561.1"
```

```
Query Match      2.5%; Score 74.8; DB 8; Length 183990;
Best Local Similarity 63.2%; Pred No. 5.8e-08;
Matches 115; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 2450 ATAAAGAGTTAAATGATGCTAGGCTCTTTGATCTTCTGAGAGTGCCACTTAGTCCAC 2509
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
82480 ATAAAGAGTAAATTGCAGGCCGGTCTTAAAGTATTATGTGTTTTCATCTAGGTCCTC 82421

QY 2510 AACTCTCAATTCGATTTTTCACACCCCTAATGTTATTCAGTGTGCCACTTAGATCTAC 2569
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
82420 AAGCTATGAATGCAATGCTGACTCTCTAATGTTATTTAAGTGTCTCATCTAGTCCCC 82361

QY 2570 AACTCTCAAAATGCATTTCTGATACCCCTAGTGTGTTTCAAGTGTCTCATCTAGGAAGA 2629
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
82360 AACTAGGAATGCAAGTCTGACTCTCTAATGTTATTTAAGTGTCTCATCTAGTCCAA 82301

QY 2630 AA 2631
Db      ||
82300 AA 82299
```

Search completed: February 16, 2005, 07:44:40
Job time : 12881 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2005, 14:41:05 ; Search time 1550 Seconds
(without alignments)
11518.671 Million cell updates/sec

Title: US-10-751-612-1

Perfect score: 3016

Sequence: 1 tctagagcataggcatgtga.....gggctatgtcaagtcctatgg 3016

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004as:*
- 13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3016	100.0	3016	13 ADR21935	Adr21935 Sugarcane
C 2	118.6	3.9	3688	2 AAZ28438	Aaz28438 Sugar can
C 3	118.6	3.9	3688	12 ADK66212	Adk66212 Sugarcane
C 4	118.6	3.9	3691	2 AAZ28433	Aaz28433 Ub19 gene
C 5	118.6	3.9	3691	12 ADK66205	Adk66205 Sugarcane
C 6	118.6	3.9	5174	2 AAZ28437	Aaz28437 Sugar can
C 7	118.6	3.9	5174	12 ADK66210	Adk66210 Sugarcane
C 8	66.8	2.2	2000	12 ADJ40707	Adj40707 Plant cDN
C 9	65.4	2.2	1334	6 ABK93956	Abk93956 cDNA enco
C 10	62	2.1	701	12 ADJ48000	Adj48000 Maize oil
C 11	61.2	2.0	290	7 ADS68705	Ads68705 Corn seed
C 12	60.4	2.0	2000	8 ADA72001	Ada72001 Rice gene
C 13	59	2.0	2010	10 ADC08563	Adc08563 Rice DNA
C 14	57.6	1.9	2000	12 ADJ41371	Adj41371 Plant cDN
C 15	53.4	1.8	2010	10 ADC08563	Adc08563 Rice DNA
C 16	53	1.8	5955	12 ADH22268	Adh22268 Rice PONG
C 17	52.6	1.7	1710	3 AAZ299113	Aaz299113 Phosphoen
C 18	52.6	1.7	1973	3 AAZ299112	Aaz299112 Phosphoen
C 19	52.6	1.7	2000	12 ADJ41371	Adj41371 Plant cDN
C 20	52.4	1.7	2000	8 ADA72420	Ada72420 Rice gene

C 21	52.4	1.7	2000	12 ADJ41383	Adj41383 Plant cDN
C 22	51.8	1.7	5955	12 ADH22268	Adh22268 Rice PONG
C 23	51.6	1.7	1150	8 ADA73250	Ada73250 Rice gene
C 24	51.6	1.7	2000	8 ADA71938	Ada71938 Rice gene
C 25	51.6	1.7	2000	8 ADA73161	Ada73161 Rice gene
C 26	51.4	1.7	2000	8 ADA71483	Ada71483 Rice gene
C 27	51.2	1.7	2000	8 ADA71894	Ada71894 Rice gene
C 28	51.2	1.7	2000	8 ADA71798	Ada71798 Rice gene
C 29	51.2	1.7	2000	8 ADA73446	Ada73446 Rice gene
C 30	50.2	1.7	76363	8 ACF30938	Acf30938 Rice cult
C 31	50.2	1.7	76363	12 ADI09997	Adi09997 Rice cult
C 32	50.2	1.7	76363	12 ADK72433	Adk72433 Rice fert
C 33	50	1.7	1334	6 ABK93956	Abk93956 cDNA enco
C 34	50	1.7	2000	12 ADJ41455	Adj41455 Plant cDN
C 35	50	1.7	76363	8 ACF30938	Acf30938 Rice cult
C 36	50	1.7	76363	12 ADI09997	Adi09997 Rice cult
C 37	50	1.7	76363	12 ADK72433	Adk72433 Rice fert
C 38	49.4	1.6	2000	8 ADA71938	Ada71938 Rice gene
C 39	48.8	1.6	2000	8 ADA72001	Ada72001 Rice gene
C 40	48.8	1.6	2000	12 ADJ40785	Adj40785 Plant cDN
C 41	48.2	1.6	2000	12 ADJ40688	Adj40688 Plant cDN
C 42	47.4	1.6	2000	8 ADA72420	Ada72420 Rice gene
C 43	47.4	1.6	2000	12 ADJ41383	Adj41383 Plant cDN
C 44	46	1.5	260	12 ADQ04630	Adq04630 Maize hom
C 45	45.8	1.5	2000	8 ADA72815	Ada72815 Rice gene

ALIGNMENTS

RESULT 1

ADR21935
ID ADR21935 standard; DNA; 3016 BP.

XX AC ADR21935;

XX DT 21-OCT-2004 (first entry)

XX DE Sugarcane O-methyl transferase (OMT) promoter DNA SeqID 1.

XX KW O-methyl transferase; OMT; promoter; stem-specific; defence-inducible;
carbon metabolism; insecticidal; pest tolerance; plant; ds.

XX OS Saccharum.

XX FH Key Location/Qualifiers

FT CAAT_signal 2661..2664

FT TATA_signal 2849..2855

FT misc_feature 3013..3015

FT /tag= a

FT /tag= b

FT /tag= c

FT /note= "Start codon (AUG) of OMT"

XX WO2004062365-A2.

XX PN 29-JUL-2004.

XX PD 05-JAN-2004; 2004WO-US000113.

XX PP 03-JAN-2003; 2003US-0437890P.

XX PR (TEXA) UNIV TEXAS A & M SYSTEM.

XX PA Mirkov TE, Damaj MB, Reddy AS, Thomas TL, Rathore KS, Emani C;

XX PI Kumpatla SP;

XX XX WPI; 2004-544018/52.

XX DR New isolated nucleic acid comprises an o-methyl transferase (OMT)

XX PT promoter and an exogenous nucleic acid, useful as a promoter for altering

XX PT carbon metabolism in a plant cell or for driving expression of

XX PT insecticidal proteins in sugarcane.

Db 1801 TAAGTAAATGCTTTGGCTTCATCACCCGGCTTAATGCTCGACAGAAAAACACGTCGCT 1860
QY 1861 AGTCAAGGTTGGCTTAACAACTGGGGTTACATGTAAACACGTTTCATGCTTAGAAA 1920
Db 1861 AGTCAAGGTTGGCTTAACAACTGGGGTTACATGTAAACACGTTTCATGCTTAGAAA 1920
QY 1921 CGGCTCGAGGATTAGATACAACTTCAATATATATCTTAGGGCCCTCCAAATATGTGTCAG 1980
Db 1921 CGGCTCGAGGATTAGATACAACTTCAATATATATCTTAGGGCCCTCCAAATATGTGTCAG 1980
QY 1981 CTCTAAATAGTGTATGTGTACCGTGGAGAGGAGGCTAAATAATATAATCTTGAG 2040
Db 1981 CTCTAAATAGTGTATGTGTACCGTGGAGAGGAGGCTAAATAATATAATCTTGAG 2040
QY 2041 CTAAGCTGAAGAGAGCTATTTTTTTTCTGCTCCCAATACATGATAGATACATATGA 2100
Db 2041 CTAAGCTGAAGAGAGCTATTTTTTTTCTGCTCCCAATACATGATAGATACATATGA 2100
QY 2101 GAGAAAAATATATGAATAAAGAACACTTTATCATGCCAGCCATACAAATATGAGATTTTCA 2160
Db 2101 GAGAAAAATATATGAATAAAGAACACTTTATCATGCCAGCCATACAAATATGAGATTTTCA 2160
QY 2161 CTAAGAGCAAACACTGACTGCTACTGTGTAAGGTGCTCTAGTTGGAGTGGTGCATCTTT 2220
Db 2161 CTAAGAGCAAACACTGACTGCTACTGTGTAAGGTGCTCTAGTTGGAGTGGTGCATCTTT 2220
QY 2221 TAGTTGTTAGTAGTAGTAGAGCTAGTTAGTGTCTTTCTTCTGCTAGTGTATGTTG 2280
Db 2221 TAGTTGTTAGTAGTAGTAGAGCTAGTTAGTGTCTTTCTTCTGCTAGTGTATGTTG 2280
QY 2281 TTTTGGCTGCCAAGTGTGTAACAACTCAAGGTAAAGTCCCATCTAATCTAAATATGATGC 2340
Db 2281 TTTTGGCTGCCAAGTGTGTAACAACTCAAGGTAAAGTCCCATCTAATCTAAATATGATGC 2340
QY 2341 CAAATAAAGATAGATTACAAAGTTAAACGAGGAAAACTCTAAATAGGATGGAAGTT 2400
Db 2341 CAAATAAAGATAGATTACAAAGTTAAACGAGGAAAACTCTAAATAGGATGGAAGTT 2400
QY 2401 TTGTAGATTAATATGTTAGTAGTAGAGTGGAGTGCACCAACCAACATATAGATTGA 2460
Db 2401 TTGTAGATTAATATGTTAGTAGTAGAGTGGAGTGCACCAACCAACATATAGATTGA 2460
QY 2461 AATGATGTTAGTGGCTCTTGTATCTCTGAGGTGCCACTTAGGTGCCACAACTCTCAA 2520
Db 2461 AATGATGTTAGTGGCTCTTGTATCTCTGAGGTGCCACTTAGGTGCCACAACTCTCAA 2520
QY 2521 TTGCAATTTTGCACCCCTTAATGTTATTCAGGTGTCCTTAGATCTACAACTCTCAA 2580
Db 2521 TTGCAATTTTGCACCCCTTAATGTTATTCAGGTGTCCTTAGATCTACAACTCTCAA 2580
QY 2581 ATGCATTTCTGATACCCCTAGTGTGTTTCAAGTGTGCTACTAGGCAAGAAAAGTTAGTA 2640
Db 2581 ATGCATTTCTGATACCCCTAGTGTGTTTCAAGTGTGCTACTAGGCAAGAAAAGTTAGTA 2640
QY 2641 ATTTGATAAGCTATGGACCAAAATAATTTATGATGATGCTCGAATAGTTGATCAT 2700
Db 2641 ATTTGATAAGCTATGGACCAAAATAATTTATGATGATGCTCGAATAGTTGATCAT 2700
QY 2701 GATGACCCCAATATAGACACTAGTTTCATGGGCTGGTTTCTTGTATAGTACTAGT 2760
Db 2701 GATGACCCCAATATAGACACTAGTTTCATGGGCTGGTTTCTTGTATAGTACTAGT 2760
QY 2761 ATAATCTTTTCAAGTGTAGTACTACTTACTTACTCGCATATATACAAATATA 2820
Db 2761 ATAATCTTTTCAAGTGTAGTACTACTTACTTACTCGCATATATACAAATATA 2820
QY 2821 GAATTCGGAAGTACTATAAAGCGGAGCTATAAATGGAGAGCTTTTGCATCATGAGGCTA 2880
Db 2821 GAATTCGGAAGTACTATAAAGCGGAGCTATAAATGGAGAGCTTTTGCATCATGAGGCTA 2880
QY 2881 TAACAACCTTGAGCAAAAAACAGAACCGGCTCGCGCATGGGCTCAGCAAGGAGCAACACACA 2940

Db 2881 TAACAACCTTGAGCAAAAAACAGAACCGGCTCGCGCATGGGCTCAGCAAGGAGCAACACACA 2940
QY 2941 AGCACTGATCAGCAGCGCGTCTGATGCTCAGTCCAGTCTCGCACCACACACCTGGGC 3000
Db 2941 AGCACTGATCAGCAGCGCGTCTGATGCTCAGTCCAGTCTCGCACCACACACCTGGGC 3000
QY 3001 TATGTCGAAGTCCATGG 3016
Db 3001 TATGTCGAAGTCCATGG 3016
RESULT 2
AAZ28438/C
ID AAZ28438 standard; DNA; 3688 BP.
XX
AC AAZ28438;
XX 23-DEC-1999 (first entry)
XX Sugar cane ubi9 gene fragment used in the construction of pubi9-GUS.
XX Ubiquitin gene; ubi9; promoter; initiation codon; intron; monocotyledon;
XX dicotyledon; disease resistance; environmental stress; sugar cane;
XX pubi4-GUS; reporter plasmid; beta-glucuronidase; GUS; ss.
XX Saccharum sp.
XX WO9946976-A1.
XX 23-SEP-1999.
XX 18-MAR-1999; 99WO-US005985.
XX 19-MAR-1998; 98US-0078768P.
XX 17-MAR-1999; 99US-00270976.
XX (UYHA-) UNIV HAWAII.
XX (USDA) US DEPT OF AGRICULTURE.
XX Albert HH, Wei H;
XX WPI; 1999-590952/50.
XX New nucleotide sequences derived from sugarcane polyubiquitin genes
XX useful in regulating expression of nucleic acid sequences in
XX monocotyledonous and dicotyledonous plants.
XX Claim 11; Fig 11; 118pp; English.
XX This sequence corresponds to nucleotides 1-3688 of the sugar cane
XX ubiquitin 9 (ubi9) gene AAZ28437. This fragment of the gene was ligated
XX to the gene encoding beta glucuronidase (GUS) in the generation of the
XX reporter plasmid pubi9-GUS and 9PI-GUS. The plasmids are used in the
XX methods of the invention, which relates to nucleotide sequences derived
XX from sugar cane polyubiquitin genes. Ubi4 and ubi9 genes are used in the
XX invention as they have linked promoter activity. The genes can be used to
XX regulate the expression of nucleic acid sequences introduced in to
XX dicotyledonous and monocotyledonous plants. The nucleic acid sequences
XX (ubi4 and ubi9) are used to create plants with commercially useful
XX phenotypes. High levels of protein expression aid the generation of
XX plants which exhibit commercially important phenotypic properties such as
XX pest and disease resistance, resistance to environmental stress and
XX improved qualities
XX Sequence 3688 BP; 844 A; 862 C; 925 G; 1055 T; 0 U; 2 Other;
SQ
Query Match 3.9%; Score 118.6; DB 2; Length 3688;
Best Local Similarity 86.0%; Pred. No. 5.4e-22;
Matches 166; Conservative 0; Mismatches 24; Indels 3; Gaps 3;
QY 1451 AGTTTCTAAGCCAGTCTCAGTGGGTTTCATCAGAGTTTCATGAGCAATTAAATAGCT 1510
Db 1914 AGTTTCTGAGCCCGTCTCAGT-GGATTTTCATCAGAGTTTCATGAGCAATTAAATAGCT 1856

QY 1511 GATGTGACACCGTATTGATGAAGAGAGAGATGATAAGAGTTTCATGCGAGTAGAGAGGT 1570
 |||||
 Db 1855 GATGTGACACCGTATTGATGAAGAGAGAGATGATAAGAGTTTCATGCGAGTAGAGAGGT 1796
 |||||
 QY 1571 TTTCATGGGATGAACACTCTT-CTTCACCTGTTTCCAAATAT-AGATGCAATTGGTAAGAGG 1628
 |||||
 Db 1795 TTTCAGAGATGAACACTCTTCTGCACTGTTTCCAAATATGGTTCGATTAAATACATG 1736
 |||||
 QY 1629 GCCATGAATCTC 1641
 |||||
 Db 1735 GCCATAAATCCC 1723
 |||||

RESULT 3

ADK66212/c
 ID ADK66212 standard; DNA; 3688 BP.

XX AC ADK66212;

XX DT 06-MAY-2004 (first entry)

XX DE Sugarcane polyubiquitin ubi9 gene promoter DNA.

XX XX Polyubiquitin; monocotyledonous; dicotyledonous; transgenic plant;

KW Phenotypic; sugarcane; plant; ds.

XX OS Saccharum sp.

XX FH Key Location/Qualifiers

FT misc_feature

1 /tag= a

/note= "Upstream of 5' UTR"

2249..2313

5'UTR /tag= b

2314..3688

intron /tag= c

FT US6686513-B1.

XX PN 03-FEB-2004.

XX PD 20-OCT-2000; 2000US-00693467.

XX PF 19-MAR-1998; 98US-0078767P.

XX PR 17-MAR-1999; 99US-00270976.

XX PR (USDA) US SEC OF AGRIC.

XX PA (UYHA-) UNIV HAWAII.

XX PA Albert HH, Wei H;

XX PI WPI; 2004-106365/11.

XX DR New transgenic plant cell comprising the sugarcane ubi9 polyubiquitin

XX PT gene promoter, useful for generating transgenic plants exhibiting

XX PT important phenotypic properties, such as pest, disease and environmental

XX PT stress resistance.

XX PT Claim 1; SEQ ID NO 10; 79pp; English.

XX PS The present invention provides nucleic acids having promoter activity.

XX CC The invention is directed to isolating nucleic acid sequences from

XX CC sugarcane polyubiquitin genes which are capable of directing constitutive

XX CC expression of a nucleic acid of interest in monocotyledonous and

XX CC dicotyledonous plants. The invention is useful in generating transgenic

XX CC plants which exhibit commercially important phenotypic properties. The

XX CC present sequence is sugarcane polyubiquitin ubi9 gene promoter DNA.

XX XX Sequence 3688 BP; 844 A; 862 C; 925 G; 1055 T; 0 U; 2 Other;

XX SQ Query Match 3.9%; Score 118.6; DB 12; Length 3688;

XX Best Local Similarity 86.0%; Pred. No. 5.4e-22;

Matches 166; Conservative 0; Mismatches 24; Indels 3; Gaps 3;
 QY 1451 AGTTTCTAAGCCAGTCTCAGTGGGTTTCATCAGAGTTTCATGACATTAATAAGCT 1510
 |||||
 Db 1914 AGTTTCTGAGGCCCTCTCACT-GGATTTCATCAGAGTTTCATGACATTAATAAGCT 1856
 |||||
 QY 1511 GATGTGACACCGTATTGATGAAGAGAGATGATAAGAGTTTCATGCGAGTAGAGAGGT 1570
 |||||
 Db 1855 GATGTGACACCGTATTGATGAAGAGAGATGATAAGAGTTTCATGCGAGTAGAGAGGT 1796
 |||||
 QY 1571 TTTCATGGGATGAACACTCTT-CTTCACCTGTTTCCAAATAT-AGATGCAATTGGTAAGAGG 1628
 |||||
 Db 1795 TTTCAGAGATGAACACTCTTCTGCACTGTTTCCAAATATGGTTCGATTAAATACATG 1736
 |||||
 QY 1629 GCCATGAATCTC 1641
 |||||
 Db 1735 GCCATAAATCCC 1723
 |||||

RESULT 4

AAZ28433/c

ID AAZ28433 standard; DNA; 3691 BP.

XX AC AAZ28433;

XX DT 23-DEC-1999 (first entry)

XX DE Ubi9 gene fragment including initiation codon and upstream sequence.

KW Ubiquitin gene; ubi9; promoter; initiation codon; intron; monocotyledon;

KW dicotyledon; disease resistance; environmental stress; sugar cane; ss.

XX OS Saccharum sp.

XX FH Key Location/Qualifiers

FT TATA_signal

2201..2208

5'UTR /tag= a

2249..2313

FT /tag= b

XX PN WO9946976-A1.

XX PD 23-SEP-1999.

XX PF 18-MAR-1999; 99WO-US005985.

XX PR 19-MAR-1998; 98US-0078768P.

XX PR 17-MAR-1999; 99US-00270976.

XX PA (UYHA-) UNIV HAWAII.

XX PA (USDA) US DEPT OF AGRICULTURE.

XX PI Albert HH, Wei H;

XX PI WPI; 1999-590952/50.

XX DR New nucleotide sequences derived from sugarcane polyubiquitin genes

XX PT useful in regulating expression of nucleic acid sequences in

XX PT monocotyledonous and dicotyledonous plants.

XX PT Claim 1; Fig 7; 118pp; English.

XX PS This sequence contains the sugar cane ubiquitin gene ubi9 initiation

XX CC codon, and the upstream sequence of the gene including the 5'

XX CC untranslated region. The gene has an intron immediately upstream of the

XX CC initiation codon. The ubi9 gene is used in the invention as it has linked

XX CC promoter activity, and the polyubiquitin genes are useful in regulating

XX CC high expression of nucleic acid sequences in monocotyledonous and

XX CC dicotyledonous plants. The nucleic acid sequences (ubi4 and ubi9) are

XX CC used to create plants with commercially useful phenotypes. High levels of

XX CC protein expression aid the generation of plants which exhibit

XX CC commercially important phenotypic properties such as pest and disease

XX CC resistance, resistance to environmental stress and improved qualities

XX SQ Sequence 3691 BP; 845 A; 862 C; 927 G; 1055 T; 0 U; 2 Other;
 Query Match 3.9%; Score 118.6; DB 2; Length 3691;
 Best Local Similarity 86.0%; Pred. No. 5.4e-22;
 Matches 166; Conservative 0; Mismatches 24; Indels 3; Gaps 3;
 QY 1451 AGTTTCTAAGCCAGCTCTCAGTGGGGTTTCATCAGAGTTTCATGACATTAAATAAGCT 1510
 DB 1914 AGTTTCTGAGCCCGTCTCAGT-GGATTTCATCAGAGTTTCATGACATTAAATAGGCT 1856
 QY 1511 GATGTGACACCGTATTGATGAAGAGAGAGATGAAGAGTTTCATGCGAGTAGAGAGT 1570
 DB 1855 GATGTGGCACCCTATTGATGAAGAGAGAGATGAAGAGTTTCATGCGAATTAGAGAGT 1796
 QY 1571 TTCATGGGGATGAACCTCTT-CTTCACCTGTTTCCAAATAT-AGATGCATTGGTAAGAGG 1628
 DB 1795 TTTACGAAGATGAACCTCTTCTGCACTGTTTCCAAATATGCGTTGCATTAAATAACATG 1736
 QY 1629 GCCATGAATCTC 1641
 DB 1735 GCCATAAATCCC 1723
 RESULT 5
 ID ADK66205/c
 XX ADK66205 standard; DNA; 3691 BP.
 AC ADK66205;
 DT 06-MAY-2004 (first entry)
 DE Sugarcane polyubiquitin ubi9 gene #1.
 KW Polyubiquitin; monocotyledonous; dicotyledonous; transgenic plant;
 KW phenotypic; sugarcane; plant; ds.
 XX Saccharum sp.
 FH Key Location/Qualifiers
 FT misc_feature 1
 FT /*tag= a
 FT /note= "Upstream of 5' UTR"
 FT 2249..2313
 FT /*tag= b
 FT intron 2314..3688
 FT /*tag= c
 XX US6686513-B1.
 PD 03-FEB-2004.
 XX 20-OCT-2000; 2000US-00693467.
 PF 19-MAR-1998; 98US-0078767P.
 PR 17-MAR-1999; 99US-00270976.
 XX (USDA) US SEC OF AGRIC.
 PA (UTHA-) UNIV HAWAII.
 XX Albert HH, Wei H;
 XX WPI; 2004-106365/11.
 XX New transgenic plant cell comprising the sugarcane ubi9 polyubiquitin
 -PT gene promoter, useful for generating transgenic plants exhibiting
 PT important phenotypic properties, such as pest, disease and environmental
 PT stress resistance.
 XX Example 1; SEQ ID NO 3; 79pp; English.
 PS The present invention provides nucleic acids having promoter activity.
 XX The invention is directed to isolating nucleic acid sequences from
 CC New nucleotide sequences derived from sugarcane polyubiquitin genes

CC sugarcane polyubiquitin genes which are capable of directing constitutive
 CC expression of a nucleic acid of interest in monocotyledonous and
 CC dicotyledonous plants. The invention is useful in generating transgenic
 CC plants which exhibit commercially important phenotypic properties. The
 CC present sequence is sugarcane polyubiquitin ubi9 gene.
 XX SQ Sequence 3691 BP; 845 A; 862 C; 927 G; 1055 T; 0 U; 2 Other;
 Query Match 3.9%; Score 118.6; DB 12; Length 3691;
 Best Local Similarity 86.0%; Pred. No. 5.4e-22;
 Matches 166; Conservative 0; Mismatches 24; Indels 3; Gaps 3;
 QY 1451 AGTTTCTAAGCCAGCTCTCAGTGGGGTTTCATCAGAGTTTCATGACATTAAATAAGCT 1510
 DB 1914 AGTTTCTGAGCCCGTCTCAGT-GGATTTCATCAGAGTTTCATGACATTAAATAGGCT 1856
 QY 1511 GATGTGACACCGTATTGATGAAGAGAGAGATGAAGAGTTTCATGCGAGTAGAGAGT 1570
 DB 1855 GATGTGGCACCCTATTGATGAAGAGAGAGATGAAGAGTTTCATGCGAATTAGAGAGT 1796
 QY 1571 TTCATGGGGATGAACCTCTT-CTTCACCTGTTTCCAAATAT-AGATGCATTGGTAAGAGG 1628
 DB 1795 TTTACGAAGATGAACCTCTTCTGCACTGTTTCCAAATATGCGTTGCATTAAATAACATG 1736
 QY 1629 GCCATGAATCTC 1641
 DB 1735 GCCATAAATCCC 1723
 RESULT 6
 ID AA228437/c
 XX AA228437 standard; DNA; 5174 BP.
 AC AA228437;
 DT 23-DEC-1999 (first entry)
 DE Sugar cane ubiquitin 9 (ubi9) gene.
 KW Ubiquitin gene; ubi4; promoter; initiation codon; intron; monocotyledon;
 KW dicotyledon; disease resistance; environmental stress; sugar cane; ss.
 XX Saccharum sp.
 FH Key Location/Qualifiers
 FT CDS 3689..3691
 FT /*tag= a
 FT /product= "ubi9"
 FT /note= "Ubiquitin 9 protein"
 FT /transl_except= (Pos:4043..4045, aa:Xaa)
 FT /transl_except= (Pos:4316..4318, aa:Xaa)
 FT /transl_except= (Pos:4499..4501, aa:Xaa)
 FT /transl_except= (Pos:4544..4547, aa:Xaa)
 FT /note= "Xaa = Unknown"
 XX WO9946976-A1.
 XX 23-SEP-1999.
 XX 18-MAR-1999; 99WO-US005985.
 XX 19-MAR-1998; 98US-0078768P.
 PR 17-MAR-1999; 99US-00270976.
 XX (UTHA-) UNIV HAWAII.
 PA (USDA) US DEPT OF AGRICULTURE.
 XX Albert HH, Wei H;
 XX WPI; 1999-590952/50.
 DR P-PSDB; AAY43047.
 XX New nucleotide sequences derived from sugarcane polyubiquitin genes

CC disease resistance (e.g. to insect infestation, fungal disease, bacterial
 CC infection, Goss' Bacterial wilt, blight, Stewart's bacterial wilt, Holcus
 CC spot, bacterial leaf blight, leaf spot, bacterial stripe and maize dwarf
 CC mosaic virus infection) and resistance to environmental stress (e.g.
 CC water stresses, pH stress, temperature stress, pollution, injury or
 CC pesticides. The present sequence is cdp cDNA sequence.
 XX
 SQ Sequence 290 BP; 58 A; 113 C; 75 G; 43 T; 0 U; 1 Other;

Query Match 2.0%; Score 61.2; DB 7; Length 290;
 Best Local Similarity 71.1%; Pred. No. 2.1e-06;
 Matches 81; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 2903 AGCCGTGGCCATGGCTCAGCAAGGAGCAGACACACAGCACTCATCAGCAGGCGGTGC 2962
 DB 10 ACCAATAGCCATGGTTCATCATGGAGCAAGCTGCACACAGCATCCGACGGCGGTGC 69

QY 2963 TGGATGCTCAGCTCAGCTTGGCCACACACCTGGCTATGTCAGTCCATGG 3016
 DB 70 TCGACGCTCAGCTCAGCTTGGCCACACAGCTTGGCTTTCATCAAGTCCATGG 123

RESULT 12
 ADA72001/c
 ID ADA72001 standard; DNA; 2000 BP.

AC ADA72001;
 DT 20-NOV-2003 (first entry)
 XX Rice gene, SEQ ID 5326.

XX Plant; bacterial infection; fungal infection; viral infection; rice;
 KW gene; db.

OS Oryza sativa.
 XX WO2003000898-A1.
 XX 03-JAN-2003.

XX 22-JUN-2001; 2001WO-IB001105.
 XX 22-JUN-2001; 2001WO-IB001105.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
 PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
 XX WPI; 2003-175290/17.

XX Identifying at least one gene involved in plant resistance or response to
 PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting plant
 PT gene expression.

XX Claim 27; SEQ ID NO 5326; 899pp; English.
 XX The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.

SQ Sequence 2000 BP; 708 A; 348 C; 340 G; 604 T; 0 U; 0 Other;
 Query Match 2.0%; Score 60.4; DB 8; Length 2000;

Best Local Similarity 62.7%; Pred. No. 1.1e-05;
 Matches 94; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 2431 AAGTCGACCACAAACATAAGAGTTAAATGATGATGGCTCTTGATCTTGTCTGG 2490
 DB 706 AAGTTCCTCTCAAAAAAAACTTTAGACTAAAGTGCATGGGTGCTCAATAAACTTTGACGG 647

QY 2491 AGGTGCCACTTAGGTCCACAAACTCTCAAAATGCAATTTTGACACCCCTAAATGTTATTCAA 2550
 DB 646 GTGTGTACCTACCTGCTCATATAAACTCTTAAATGTTATTTTGGATCCGTCGTTATCTC 587

QY 2551 GTGTGCCACTTAGATCTACAAACTCTCAAA 2580
 DB 586 GGGTGTCTATGCTGGGTCTAAATGGGCTCAA 557

RESULT 13
 ADC08563/c
 ID ADC08563 standard; DNA; 2010 BP.

AC ADC08563;

XX 18-DEC-2003 (first entry)

XX Rice DNA sequence Seq ID868 related to grain filling.

XX plant biotechnology; carbohydrate synthesis; carbohydrate metabolism;
 KW carbohydrate degradation; carbohydrate; plant grain; grain filling; corn;
 KW tomato; banana; canola; cotton; peanut; sorghum; tobacco; sugarbeet;
 KW wheat; rice; protein; oil; starch; fibre; moisture content; cereal grain;
 KW gene; db; plant.

OS Oryza sativa.

XX WO2003000905-A2.

XX 03-JAN-2003.

XX 21-JUN-2002; 2002WO-IB002450.

XX 22-JUN-2001; 2001US-0300112P.

XX 26-SEP-2001; 2001US-0325277P.

XX 20-DEC-2001; 2001US-0342327P.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Zhu T, Cheng W, Briggs S, Cooper B, Goff SA, Moughamer T;
 PI Glazebrook J, Katagiri F, Kreps J, Provart N, Ricke D;
 XX WPI; 2003-229341/22.

XX New plant genes encoding polypeptides having an activity involved in or
 PT associated with the synthesis, metabolism or degradation of carbohydrates
 PT in the plant grain useful in generating plants having improved
 PT nutritional properties.

XX Disclosure; SEQ ID NO 868; 130pp; English.

XX This invention, in the area of plant biotechnology, relates to novel
 CC polynucleotides comprising a nucleotide sequence encoding a protein which
 CC is involved in or associated with the synthesis, metabolism or
 CC degradation of carbohydrates in the plant grain and the expression of
 CC which is up-regulated during grain filling. The plant is selected from
 CC corn, tomato, banana, canola, cotton, peanut, sorghum, tobacco,
 CC sugarbeet, wheat, and rice. The invention may be useful for the
 CC improvement of protein, oil, starch, fibre and moisture content of the
 CC cereal grains. In addition, carbohydrate levels may be modified to a more
 CC desirable level using the present invention. The present sequence is a
 CC DNA sequence of a rice gene promoter. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/publishedpct_sequences.

SQ Sequence 2010 BP; 563 A; 472 C; 389 G; 586 T; 0 U; 0 Other;

Query Match 2.0%; Score 59; DB 10; Length 2010;
 Best Local Similarity 58.0%; Pred. No. 2.7e-05;
 Matches 141; Conservative 0; Mismatches 100; Indels 2; Gaps 2;

QY 2434 TCACACACACCAACATAAAGAGTTAAATGCATGTAGGCTCTTGATCTTCTGGAGG 2493
 DB 918 TCTTCATATCGCAGTAAATAGTAAAGTGACGGCGGTCTTAACTTTAGGGGTG 859

QY 2494 TGCCACTTAGGTCACAAACTCTCAAAATGCAATTTTGCACCCCTAAATTTTCAAGTG 2553
 DB 858 TGTCAATTAGGTCCCTAAACTCTCAAAATGCATATCCAAAGTCCAGAACTTGTATAGTG 799

QY 2554 TGCCACTTAGATCACAACCTCTCAAAATGCAATTTTGCATACCTAGTGTGTTCAAGTG 2613
 DB 798 TATCATCTAGGT-TCCAAATCGCCACACCCCTTCAGGATCTATGTGCAATTGATGG 740

QY 2614 TGTCACTTAGGCAAGAAAGTTAGATAATTTTGA-TAAGCTATGGACCAAAATTAATTA 2672
 DB 739 CATGCCACATGGACATGACGTGGTATATTTTGAAGTGCACAAATGGACCAATTAATTT 680

QY 2673 TGT 2675
 DB 679 TTT 677

RESULT 14
 ADJ41371
 ID ADJ41371 standard; cDNA; 2000 BP.
 AC ADJ41371;
 DT 06-MAY-2004 (first entry)
 XX Plant cDNA #2371.
 DE
 KW plant; gene; ss; transcription; plant genome augmentation; cereal;
 KW soybean; alfalfa; sunflower; canola; cotton; peanut; tobacco; sugar beet;
 KW maize; barley; sorghum; rice; wheat; crop plant; insecticide resistance;
 KW stress tolerance; salt tolerance; cold tolerance; drought tolerance;
 KW plant nutrition; apical dominance; dwarfism; early flowering; antiviral;
 KW antifungal.
 XX Eukaryota.
 OS
 XX US2004016025-A1.
 PN
 XX 22-JAN-2004.
 PD
 XX 26-SEP-2002; 2002US-00260238.
 PF
 XX 26-SEP-2001; 2001US-0325277P.
 PR 26-SEP-2001; 2001US-0325448P.
 PR 04-APR-2002; 2002US-0370620P.
 XX (BUDW/) BUDWORTH P.
 PA (MOUG/) MOUGHAMER T.
 PA (BRIG/) BRIGGS S F.
 PA (COOP/) COOPER B.
 PA (GLAZ/) GLAZEBOOK J.
 PA (GOLF/) GOLF S A.
 PA (KATA/) KATAGIRI F.
 PA (KREP/) KREPS J.
 PA (PROV/) PROVART N.
 PA (RICK/) RIQUE D.
 PA (ZHUT/) ZHU T.
 XX Budworth P, Moughamer T, Briggs SP, Cooper B, Glazebrook J;
 PI Goff SA, Katagiri F, Kreps J, Provart N, Ricke D, Zhu T;
 XX WPI; 2004-190374/18.
 DR
 XX

PT New rice promoter, useful for manipulating crop plants to alter or
 PT improve phenotypic characteristics, e.g. produce large quantities of oil
 PT or proteins, resistance to insecticides, virus or fungi, stress tolerance
 PT or high nutritional value.
 XX Claim 1; SEQ ID NO 2371; 230pp; English.
 PS
 XX The invention relates to plant nucleotide sequences that direct seed-,
 CC leaf- and/or stem-, panicle-, root- or pollen-specific or -preferential
 CC or constitutive transcription of an operatively linked nucleic acid
 CC segment. The invention also relates to a method for augmenting a plant
 CC genome and a method of identifying a gene, where its expression is
 CC altered in the seed, leaf, stem, panicle, pollen, root or is constitutive
 CC in a plant cell. The plant is a cereal, e.g. soybean, alfalfa, sunflower,
 CC canola, cotton, peanut, tobacco or sugar beet, preferably maize, barley,
 CC sorghum, rice or wheat. The polynucleotides and the polypeptides they
 CC encode are useful for manipulating crop plants to alter or improve
 CC phenotypic characteristics, to produce large quantities of oil or
 CC proteins, to incur resistance to insecticides, viruses or fungi, and to
 CC incur stress tolerance (e.g. salt, cold or drought) to ensure the plants
 CC have a high nutritional value with reduced apical dominance or dwarfism,
 CC early flowering or altered metabolic pathways. This sequence represents a
 CC plant nucleic acid of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification but was obtained in
 CC electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.
 XX Sequence 2000 BP; 662 A; 359 C; 368 G; 609 T; 0 U; 2 Other;
 SQ

Query Match 1.9%; Score 57.6; DB 12; Length 2000;
 Best Local Similarity 56.2%; Pred. No. 6.7e-05;
 Matches 108; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 2448 ACATTAAGAGTTAAATGCATGTAGGCTCTTGATCTTCTGGAGTCCACTTAGGTCC 2507
 DB 885 AAAAAATAGATAAAGTCCATCACCGTCCCTCAAACTTGTACTGTGTCTACCGGTCC 944

QY 2508 ACAAACTCTCAAAATGCAATTTTGCACACCCCTAAATGTTATTCAGTGTCCTAGATCT 2567
 DB 945 CTAACCTCGCAATTCACCGTTCAGTCTCTCAAACTTGTCAATGTTCTATCCGGTCC 1004

QY 2568 ACAAACTCTCAAAATGCAATTTTGCATACCTTGTGTTTCAAGTGTGTCTACTTAGGCAA 2627
 DB 1005 CTAAACTTACAGATCACTCGTTTAGTCTCTCAAACTTGTTCAGTTGTGTCTACCCCGGTCT 1064

QY 2628 GAAAGTTAGAT 2639
 DB 1065 CTAACCTTGGAT 1076

RESULT 15
 ADC08563
 ID ADC08563 standard; DNA; 2010 BP.
 XX
 AC ADC08563;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE Rice DNA sequence Seq ID868 related to grain filling.
 XX
 KW plant biotechnology; carbohydrate synthesis; carbohydrate metabolism;
 KW carbohydrate degradation; carbohydrate; plant grain; grain filling; corn;
 KW tomato; banana; canola; cotton; peanut; sorghum; tobacco; sugarbeet;
 KW wheat; rice; protein; oil; starch; fibre; moisture content; cereal grain;
 KW gene; ds; plant.
 XX
 OS Oryza sativa.
 XX
 XX WO2003000905-A2.
 PN
 XX 03-JAN-2003.
 PD
 XX 21-JUN-2002; 2002MO-IB002450.
 PF
 XX

PR 22-JUN-2001; 2001US-0300112P.
PR 26-SEP-2001; 2001US-0325277P.
PR 20-DEC-2001; 2001US-0342327P.
XX
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
PA
XX Zhu T, Cheng W, Briggs S, Cooper B, Goff SA, Moughamer T;
PI Glazebrook J, Katagiri F, Kreps J, Provart N, Ricke D;
XX WPI; 2003-229341/22.
DR
XX
PT New plant genes encoding polypeptides having an activity involved in or
PT associated with the synthesis, metabolism or degradation of carbohydrates
PT in the plant grain useful in generating plants having improved
PT nutritional properties.
XX
PS Disclosure; SEQ ID NO 868; 130pp; English.
XX
CC This invention, in the area of plant biotechnology, relates to novel
CC polynucleotides comprising a nucleotide sequence encoding a protein which
CC is involved in or associated with the synthesis, metabolism or
CC degradation of carbohydrates in the plant grain and the expression of
CC which is up-regulated during grain filling. The plant is selected from
CC corn, tomato, banana, canola, cotton, peanut, sorghum, tobacco,
CC sugarbeet, wheat, and rice. The invention may be useful for the
CC improvement of protein, oil, starch, fibre and moisture content of the
CC cereal grains. In addition, carbohydrate levels may be modified to a more
CC desirable level using the present invention. The present sequence is a
CC DNA sequence of a rice gene promoter. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/publishedpct_sequences.
XX
SQ Sequence 2010 BP; 563 A; 472 C; 389 G; 586 T; 0 U; 0 Other;

Query Match 1.8%; Score 53.4; DB 10; Length 2010;
Best Local Similarity 65.5%; Pred. No. 0.001;
Matches 78; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 2455 GAGTTAATGCATGGTAGGCTTGTATCTTGTCTGAGGTGCCACTTAGGTCCAAACT 2514
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
367 GAGTAAAGTGCATGGCGGTCTTAATCTTGTAGGGTTGTGCATATAGTCCCTAAACT 426
QY 2515 CTCAAATTGCATTTTGACACCCCTAATGTTATTCAGTGTGCCACTTAGATCTCAAAAC 2573
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
427 CCCAAATGCATATCCAGGTCCGAACTTATCAAAAGTGATCATCTAGTCCCAATC 485

Search completed: February 16, 2005, 04:09:58
Job time : 1555 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 16, 2005, 02:54:17 ; Search time 490 Seconds

(without alignments)
10071.450 Million cell updates/sec

Title: US-10-751-612-1

Perfect score: 3016

Sequence: 1 tctagagcataggcattgta.....gggctatgcaagtcctcagg 3016

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	118.6	3.9	3688	4	US-09-866-153-10
C 2	118.6	3.9	3688	4	US-09-693-467A-10
C 3	118.6	3.9	3688	4	US-09-270-976-10
C 4	118.6	3.9	3691	4	US-09-866-153-3
C 5	118.6	3.9	3691	4	US-09-693-467A-3
C 6	118.6	3.9	3691	4	US-09-270-976-3
C 7	118.6	3.9	5174	4	US-09-866-153-8
C 8	118.6	3.9	5174	4	US-09-693-467A-8
C 9	118.6	3.9	5174	4	US-09-270-976-8
C 10	49.4	1.6	1141	4	US-09-806-708B-22
C 11	46	1.5	1141	4	US-09-806-708B-22
C 12	44	1.5	832	4	US-09-621-976-2813
C 13	41.2	1.4	832	4	US-09-621-976-2813
C 14	40.8	1.4	601	4	US-09-949-016-150744
C 15	40.8	1.4	294836	4	US-09-949-016-150744
C 16	40.2	1.3	4467	2	US-08-565-907A-1
C 17	40.2	1.3	4467	2	US-08-910-551B-1
C 18	40.2	1.3	4467	2	US-08-909-425A-1
C 19	39.4	1.3	1677	2	US-08-684-101-1
C 20	39.4	1.3	1677	3	US-09-205-814-1
C 21	39.4	1.3	7218	1	US-08-232-463-14
C 22	39.2	1.3	636	3	US-08-998-416-1137
C 23	39	1.3	505	4	US-09-621-976-15639
C 24	39	1.3	601	4	US-09-949-016-107633
C 25	39	1.3	837	3	US-08-998-416-288
C 26	39	1.3	390890	4	US-09-949-016-14720
C 27	38.6	1.3	4253	4	US-08-956-171E-118

28	38.6	1.3	4253	4	US-08-781-986A-118	Sequence 118, Appl
29	38.4	1.3	2755	3	US-08-749-522-2	Sequence 2, Appli
C 30	38.4	1.3	19480	4	US-09-949-016-14095	Sequence 14095, A
C 31	38.4	1.3	387902	4	US-09-949-016-14543	Sequence 14543, A
C 32	38.4	1.3	421883	4	US-09-949-016-12557	Sequence 12557, A
C 33	38	1.3	17633	4	US-09-949-016-13137	Sequence 13137, A
C 34	37.6	1.2	1113	4	US-09-248-796A-9644	Sequence 9644, Ap
C 35	37.4	1.2	2287	3	US-08-845-258-8	Sequence 8, Appli
C 36	37.4	1.2	2287	3	US-08-990-571-8	Sequence 8, Appli
C 37	37.4	1.2	2287	3	US-08-723-142A-8	Sequence 8, Appli
C 38	37.4	1.2	2287	3	US-09-528-784A-8	Sequence 8, Appli
C 39	37.4	1.2	2287	4	US-09-569-098A-8	Sequence 8, Appli
C 40	37.4	1.2	38575	4	US-09-949-016-12876	Sequence 12876, A
C 41	37.4	1.2	38575	4	US-09-949-016-13871	Sequence 13871, A
C 42	37.2	1.2	2430	3	US-08-845-258-3	Sequence 3, Appli
C 43	37.2	1.2	2430	3	US-08-845-258-40	Sequence 40, Appl
C 44	37.2	1.2	2430	3	US-08-990-571-3	Sequence 3, Appli
C 45	37.2	1.2	2430	3	US-08-990-571-40	Sequence 40, Appl

ALIGNMENTS

RESULT 1
US-09-866-153-10/c
; Sequence 10, Application US/09866153
; Patent No. 6638766
; GENERAL INFORMATION:
; APPLICANT: Albert, Henrik H.
; APPLICANT: Wei, Hairong
; TITLE OF INVENTION: PLANT PROMOTER SEQUENCES AND METHODS OF USE THEREOF
; FILE REFERENCE: US-03648
; CURRENT APPLICATION NUMBER: US/09/866,153
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/270,976
; PRIOR FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 3688
; TYPE: DNA
; ORGANISM: Saccharum Hybrid Cultivar 32-8560
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9)
; OTHER INFORMATION: The "n" at position 9 is any nucleotide.
US-09-866-153-10

Query Match	3.9%	Score 118.6;	DB 4;	Length 3688;
Best Local Similarity	86.0%	Pred. No. 1.3e-25;		
Matches 166;	Conservative 0;	Mismatches 24;	Indels 3;	Gaps 3;
QY	1451	AGTTTCTTAAGCCAGTCTCAGTGGGTTTCATCAGAGTTTCATGAGCAATTAAGCT	1510	
Db	1914	AGTTTCTTGAAGCCCGTCTCAGT-GGATTTCATCAGAGTTTCATGAGCAATTAAGCT	1856	
QY	1511	GATGTGACACCGTATTGATGAAGAGAGATGAAGAGTTTCATGCCAGTAGAGAGCT	1570	
Db	1855	GATGTGACACCGTATTGATGAAGAGAGATGAAGAGTTTCATGCCAGTAGAGAGCT	1796	
QY	1571	TTTCATGGGATGAACCTCTT-CCTTCACTGTTTCCAAAATAT-AGATGCATTGGTAAAGG	1628	
Db	1795	TTTACGAAGATGAACCTCTTCTCTGCACCTGTTTCCAAATATGGGTTGCTAATTAACATG	1736	
QY	1629	GCCATGAATCTC	1641	
Db	1735	GCCATAAATCCC	1723	

RESULT 2
US-09-693-467A-10/c
; Sequence 10, Application US/09693467A
; Patent No. 6686513

GENERAL INFORMATION:
; APPLICANT: Albert, Henrik H.
; APPLICANT: Wei, Hairong
; TITLE OF INVENTION: PLANT PROMOTER SEQUENCES AND METHODS OF USE THEREOF
; FILE REFERENCE: US-04331
; CURRENT APPLICATION NUMBER: US/09/693,467A
; CURRENT FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/270,976
; PRIOR FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 3688
; TYPE: DNA
; ORGANISM: Saccharum Hybrid Cultivar 32-8560
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9)
; OTHER INFORMATION: The "n" at position 9 is any nucleotide.
US-09-693-467A-10
Query Match 3.9%; Score 118.6; DB 4; Length 3688;
Best Local Similarity 86.0%; Pred. No. 1.3e-25;
Matches 166; Conservative 0; Mismatches 24; Indels 3; Gaps 3;
QY 1451 AGTTTCTAAGGCCAGTCTCAGTGGGTTTCATCAGAGTTTCATGACATTAATAAGCT 1510
DB 1914 AGTTTCTCAGGCCCTCTCAGT-GGATTTTCATCAGAGTTTCATGACATTAATAAGCT 1856
QY 1511 GATGTGACACCGTATTGATGAAGAGAGATGATGAAGAGTTTCATGCGAGTAGAGAGT 1570
DB 1855 GATGTGACACCGTATTGATGAAGAGAGATGATGAAGAGTTTCATGCGAGTAGAGAGT 1796
QY 1571 TTTCATGGGATGAAACTCTT-CTTCACTGTTTCCAAATAT-AGATGCATTGGTAAGAGG 1628
DB 1795 TTTCAGAGATGAAACTCTTCTCGCACTGTTTCCAAATATGGGTTGCAATTAATAACATG 1736
QY 1629 GCCATGAATCTC 1641
DB 1735 GCCATAAATCCC 1723

RESULT 3
US-09-270-976-10/c
; Sequence 10, Application US/09270976A
; Patent No. 6706948
; GENERAL INFORMATION:
; APPLICANT: Albert, Henrik H.
; APPLICANT: Wei, Hairong
; TITLE OF INVENTION: PLANT PROMOTER SEQUENCES AND METHODS OF USE THEREOF
; FILE REFERENCE: US-03648
; CURRENT APPLICATION NUMBER: US/09/270,976A
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 3688
; TYPE: DNA
; ORGANISM: Saccharum Hybrid Cultivar 32-8560
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9)
; OTHER INFORMATION: The "n" at position 9 is any nucleotide.
US-09-270-976-10
Query Match 3.9%; Score 118.6; DB 4; Length 3688;
Best Local Similarity 86.0%; Pred. No. 1.3e-25;
Matches 166; Conservative 0; Mismatches 24; Indels 3; Gaps 3;
QY 1451 AGTTTCTAAGGCCAGTCTCAGTGGGTTTCATCAGAGTTTCATGACATTAATAAGCT 1510
DB 1914 AGTTTCTCAGGCCCTCTCAGT-GGATTTTCATCAGAGTTTCATGACATTAATAAGCT 1856

QY 1511 GATGTGACACCGTATTGATGAAGAGAGATGATGAAGAGTTTCATGCGAGTAGAGAGT 1570
DB 1855 GATGTGACACCGTATTGATGAAGAGAGATGATGAAGAGTTTCATGCGAGTAGAGAGT 1796
QY 1571 TTTCATGGGATGAAACTCTT-CTTCACTGTTTCCAAATAT-AGATGCATTGGTAAGAGG 1628
DB 1795 TTTCAGAGATGAAACTCTTCTCGCACTGTTTCCAAATATGGGTTGCAATTAATAACATG 1736
QY 1629 GCCATGAATCTC 1641
DB 1735 GCCATAAATCCC 1723
RESULT 4
US-09-866-153-3/c
; Sequence 3, Application US/09866153
; Patent No. 6638766
; GENERAL INFORMATION:
; APPLICANT: Albert, Henrik H.
; APPLICANT: Wei, Hairong
; TITLE OF INVENTION: PLANT PROMOTER SEQUENCES AND METHODS OF USE THEREOF
; FILE REFERENCE: US-03648
; CURRENT APPLICATION NUMBER: US/09/866,153
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/270,976
; PRIOR FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 3691
; TYPE: DNA
; ORGANISM: Saccharum Hybrid Cultivar H32-8560
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9)
; OTHER INFORMATION: The "n" at position 9 is any nucleotide.
; NAME/KEY: misc_feature
; LOCATION: (3613)
; OTHER INFORMATION: The "n" at position 3613 is any nucleotide.
US-09-866-153-3
Query Match 3.9%; Score 118.6; DB 4; Length 3691;
Best Local Similarity 86.0%; Pred. No. 1.3e-25;
Matches 166; Conservative 0; Mismatches 24; Indels 3; Gaps 3;
QY 1451 AGTTTCTAAGGCCAGTCTCAGTGGGTTTCATCAGAGTTTCATGACATTAATAAGCT 1510
DB 1914 AGTTTCTCAGGCCCTCTCAGT-GGATTTTCATCAGAGTTTCATGACATTAATAAGCT 1856
QY 1511 GATGTGACACCGTATTGATGAAGAGAGATGATGAAGAGTTTCATGCGAGTAGAGAGT 1570
DB 1855 GATGTGACACCGTATTGATGAAGAGAGATGATGAAGAGTTTCATGCGAGTAGAGAGT 1796
QY 1571 TTTCATGGGATGAAACTCTT-CTTCACTGTTTCCAAATAT-AGATGCATTGGTAAGAGG 1628
DB 1795 TTTCAGAGATGAAACTCTTCTCGCACTGTTTCCAAATATGGGTTGCAATTAATAACATG 1736
QY 1629 GCCATGAATCTC 1641
DB 1735 GCCATAAATCCC 1723

RESULT 5
US-09-693-467A-3/c
; Sequence 3, Application US/09693467A
; Patent No. 6686513
; GENERAL INFORMATION:
; APPLICANT: Albert, Henrik H.
; APPLICANT: Wei, Hairong
; TITLE OF INVENTION: PLANT PROMOTER SEQUENCES AND METHODS OF USE THEREOF
; FILE REFERENCE: US-04331
; CURRENT APPLICATION NUMBER: US/09/693,467A
; CURRENT FILING DATE: 2000-10-20
Query Match 3.9%; Score 118.6; DB 4; Length 3691;
Best Local Similarity 86.0%; Pred. No. 1.3e-25;
Matches 166; Conservative 0; Mismatches 24; Indels 3; Gaps 3;
QY 1451 AGTTTCTAAGGCCAGTCTCAGTGGGTTTCATCAGAGTTTCATGACATTAATAAGCT 1510
DB 1914 AGTTTCTCAGGCCCTCTCAGT-GGATTTTCATCAGAGTTTCATGACATTAATAAGCT 1856
QY 1511 GATGTGACACCGTATTGATGAAGAGAGATGATGAAGAGTTTCATGCGAGTAGAGAGT 1570
DB 1855 GATGTGACACCGTATTGATGAAGAGAGATGATGAAGAGTTTCATGCGAGTAGAGAGT 1796
QY 1571 TTTCATGGGATGAAACTCTT-CTTCACTGTTTCCAAATAT-AGATGCATTGGTAAGAGG 1628
DB 1795 TTTCAGAGATGAAACTCTTCTCGCACTGTTTCCAAATATGGGTTGCAATTAATAACATG 1736
QY 1629 GCCATGAATCTC 1641
DB 1735 GCCATAAATCCC 1723

```

Qy 1511 GATGTGACACCGTATTGATGAAGAGAGAGATGATAAGAGTTTCATGCGAGTAGAGAGGT 1570
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1855 GATGTGGCACCGTATTGATGAAGAGAGAGATGATAAGAGTTTCATGCGAATTAGAGAGGT 1796
Qy 1571 TTCATGGGATGAAACTCTT-CTTCACTGTTTCCAAAATAT-AGATGCATTGGTAAGAGG 1628
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1795 TTTTACGAAGATGAAACTCTTCTCGCACTGTTTCCAAAATATGGGTTTGCATTAATAACATG 1736
Qy 1629 GCCATGAAATCTC 1641
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1735 GCCATTAATCCC 1723
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 7
US-09-866-153-8/c
; Sequence 8, Application US/09866153
; Patent No. 6638766
; GENERAL INFORMATION:
; APPLICANT: Albert, Henrik H.
; APPLICANT: Wei, Hairong
; TITLE OF INVENTION: PLANT PROMOTER SEQUENCES AND METHODS OF USE THEREOF
; FILE REFERENCE: UH-03648
; CURRENT APPLICATION NUMBER: US/09/866,153
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/270,976
; PRIOR FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 5174
; TYPE: DNA
; ORGANISM: Saccharum Hybrid Cultivar H32-8560
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9)
; OTHER INFORMATION: The "n" at position 9 is any nucleotide.
; NAME/KEY: misc_feature
; LOCATION: (4318)
; OTHER INFORMATION: The "n" at position 4318 is any nucleotide.
; NAME/KEY: misc_feature
; LOCATION: (4546)
; OTHER INFORMATION: The "n" at position 4546 is any nucleotide.
; NAME/KEY: misc_feature
; LOCATION: (4890)
; OTHER INFORMATION: The "n" at position 4890 is any nucleotide.
; NAME/KEY: misc_feature
; LOCATION: (5117)
; OTHER INFORMATION: The "n" at position 5117 is any nucleotide.
; NAME/KEY: misc_feature
; LOCATION: (5125)
; OTHER INFORMATION: The "n" at position 5125 is any nucleotide.
; NAME/KEY: misc_feature
; LOCATION: (5150)
; OTHER INFORMATION: The "n" at position 5150 is any nucleotide.
US-09-866-153-8

Query Match 3.9%; Score 118.6; DB 4; Length 5174;
Best Local Similarity 86.0%; Pred. No. 1.6e-25;
Matches 166; Conservative 0; Mismatches 24; Indels 3; Gaps 3;

Qy 1451 AGGTTTCTAAGGCCAGTCTCAGTGGGGTTTCATCAGAGTTTCATGGACATTAATAAGCT 1510
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1914 AGTTTCTGAGGCCGCTCAGT-GGATTTTCATCAGAGTTTCATGGACATTAATAAGCT 1856
Qy 1511 GATGTGACACCGTATTGATGAAGAGAGAGATGATAAGAGTTTCATGCGAGTAGAGAGGT 1570
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1855 GATGTGGCACCGTATTGATGAAGAGAGAGATGATAAGAGTTTCATGCGAATTAGAGAGGT 1796
Qy 1571 TTCATGGGATGAAACTCTT-CTTCACTGTTTCCAAAATAT-AGATGCATTGGTAAGAGG 1628
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1795 TTTTACGAAGATGAAACTCTTCTCGCACTGTTTCCAAAATATGGGTTTGCATTAATAACATG 1736

```

QY 1629 GCATGAATCTC 1641
|||||
Db 1735 GCATAAATCCC 1723

RESULT 8
US-09-693-467A-8/c
; Sequence 8, Application US/09693467A
; Patent No. 686513
; GENERAL INFORMATION:
; APPLICANT: Albert, Henrik H.
; APPLICANT: Wei, Hairong
; TITLE OF INVENTION: PLANT PROMOTER SEQUENCES AND METHODS OF USE THEREOF
; FILE REFERENCE: UH-04331
; CURRENT APPLICATION NUMBER: US/09/693,467A
; CURRENT FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/270,976
; PRIOR FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 8
; LENGTH: 5174
; TYPE: DNA
; ORGANISM: Saccharum Hybrid Cultivar H32-8560
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9)
; OTHER INFORMATION: The "n" at position 9 is any nucleotide.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4318)
; OTHER INFORMATION: The "n" at position 4318 is any nucleotide.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4546)
; OTHER INFORMATION: The "n" at position 4546 is any nucleotide.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4890)
; OTHER INFORMATION: The "n" at position 4890 is any nucleotide.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5117)
; OTHER INFORMATION: The "n" at position 5117 is any nucleotide.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5125)
; OTHER INFORMATION: The "n" at position 5125 is any nucleotide.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5150)
; OTHER INFORMATION: The "n" at position 5150 is any nucleotide.
US-09-693-467A-8

Query Match 3.9%; Score 118.6; DB 4; Length 5174;
Best Local Similarity 86.0%; Pred. No. 1.6e-25;
Matches 166; Conservative 0; Mismatches 24; Indels 3; Gaps 3;
QY 1451 AGGTTTCTAAGCCAGTCTCAGTGGGGTTTCATCAGAGTTTCATCGACATTAAATAGCT 1510
|||
Db 1914 AGTTTCTGAGCCCGTCTCAGT-GGATTCATCAGAGTTTCATCGACATTAAATAGCT 1856
|||
QY 1511 GATGTGACACCGTATTGATGAAGAGAGAGATGAAGAGTTTCATCGAGTAGAGAGT 1570
|||
Db 1855 GATGTGGCACCGTATTGATGAAGAGAGAGATGAAGAGTTTCATCGAATTAGAGAGT 1796
|||
QY 1571 TTTCATGGGATGAACCTCTT-CTTCACCTGTTTCCAAATAT-AGATGCATTGGTAAGAGG 1628
|||
Db 1795 TTTCAGAGATGAACCTCTCTCTGCACTGTTTCCAAATATGGGTTCATTAAATACATG 1736
|||
QY 1629 GCATGAATCTC 1641
|||||
Db 1735 GCATAAATCCC 1723

RESULT 9
US-09-270-976-8/c
; Sequence 8, Application US/09270976A
; Patent No. 6706948
; GENERAL INFORMATION:
; APPLICANT: Albert, Henrik H.
; APPLICANT: Wei, Hairong
; TITLE OF INVENTION: PLANT PROMOTER SEQUENCES AND METHODS OF USE THEREOF
; FILE REFERENCE: UH-03648
; CURRENT APPLICATION NUMBER: US/09/270,976A
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 8
; LENGTH: 5174
; TYPE: DNA
; ORGANISM: Saccharum Hybrid Cultivar H32-8560
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9)
; OTHER INFORMATION: The "n" at position 9 is any nucleotide.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4318)
; OTHER INFORMATION: The "n" at position 4318 is any nucleotide.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4546)
; OTHER INFORMATION: The "n" at position 4546 is any nucleotide.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4890)
; OTHER INFORMATION: The "n" at position 4890 is any nucleotide.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5117)
; OTHER INFORMATION: The "n" at position 5117 is any nucleotide.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5125)
; OTHER INFORMATION: The "n" at position 5125 is any nucleotide.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5150)
; OTHER INFORMATION: The "n" at position 5150 is any nucleotide.
US-09-270-976-8

Query Match 3.9%; Score 118.6; DB 4; Length 5174;
Best Local Similarity 86.0%; Pred. No. 1.6e-25;
Matches 166; Conservative 0; Mismatches 24; Indels 3; Gaps 3;
QY 1451 AGGTTTCTAAGCCAGTCTCAGTGGGGTTTCATCAGAGTTTCATCGACATTAAATAGCT 1510
|||
Db 1914 AGTTTCTGAGCCCGTCTCAGT-GGATTCATCAGAGTTTCATCGACATTAAATAGCT 1856
|||
QY 1511 GATGTGACACCGTATTGATGAAGAGAGAGATGAAGAGTTTCATCGAGTAGAGAGT 1570
|||
Db 1855 GATGTGGCACCGTATTGATGAAGAGAGAGATGAAGAGTTTCATCGAATTAGAGAGT 1796
|||
QY 1571 TTTCATGGGATGAACCTCTT-CTTCACCTGTTTCCAAATAT-AGATGCATTGGTAAGAGG 1628
|||
Db 1795 TTTCAGAGATGAACCTCTCTCTGCACTGTTTCCAAATATGGGTTCATTAAATACATG 1736
|||
QY 1629 GCATGAATCTC 1641
|||||
Db 1735 GCATAAATCCC 1723

RESULT 10
US-09-806-708B-22
; Sequence 22, Application US/09806708B

```

; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAEI promoters
US-09-806-708B-22

Query Match
Best Local Similarity 1.6%; Score 49.4; DB 4; Length 1141;
Matches 81; Conservative 311; Mismatches 432; Indels 6; Gaps 1;

QY 2023 TAAAAATATATCTTGGAGCTACGCTGAAGAGAGAGCTATTTTCTGCTCCCAATAC 2082
DB 130 WRAMBDTVDDHVYVYANNAATWTCMDKDDKRRWKKNNATGDDDDTKYHMWNNGC 189
QY 2083 ATGATAGATACATATGAGAGAAAAATATATGAATAAGAACACATTTTACATGCCAGCA 2142
DB 190 BTVTWVRVYKTRDWSBKRMNYGMBWKNWSYDVTVYVWVDDMDCKRKVRVRWVRTRGRMR 249
QY 2143 TACAATATGAGATTTCACTTAAGAGCCACACCTGCTGCTACTGCTTAAGGTGCTCCTAG 2202
DB 250 NYVMAWBAHRRRYNNGWTBAMAYRRWTMNNNNNAKAMCRKRYWGNWRABVNSTCTTW 309
QY 2203 TTGGAGTGTGCTCATCTTTTAGTTGTTAGTAGTGAAGACCTAGTTTAGTGTCTTTCTT 2262
DB 310 KSKTTKVRTSCWANNCRAGDANKDHWKWSAMGVYNNNNNNWTKKARHBABWV 369
QY 2263 GTCTAGGTTATGTTGTTGTTGCTGCAAGTGTGAACAACCTCAAGTAAGTCCCAT 2322
DB 370 WSAWKWKHANAHAHYSRKWTBYKRTWVNNNGTTWTKRMWAWYWKMDMDWBGTYNNNN 429
QY 2323 CTAATCTTAAATGATGCCAAATAAGATAGATTACAAGATTAAACGAGCGGAAAACCT 2382
DB 430 NGRTYTGWTKGNKWYTYKWKANNCKRWARDHKTCTHNTTWWOKTY-----WNNCY 483
QY 2383 AAAATAGGATGAAAGTTTGTAGAGTAATAATTGATGAGTGGCGAAGTCGACCA 2442
DB 484 WKSMTNGKSHRBAAYVYVWVWVRYAHANNNDWYWKACTWYKYBVCSKWNWYAAW 543
QY 2443 ACCAAACATAAAGAGCTAAATGATGAGTGTGCTTGAATCTTGTCTGAGGAGTGCACTTA 2502
DB 544 YTKSSWNTYSRYRYWKTNNWRSWSDTRSMGRANNYARABHYGYKWNTRWBSHTWBH 603
QY 2503 GGTCCACAACTCTCAAAATGCAATTTTGGACACCCTAATGTTATTCAAGTGTGCCACTTA 2562
DB 604 RAGAAHYWMBWBYBAKCHCKAWYKAKYAGAGGNNNNNNNNNNNNNNNNNNATCAARDYY 663
QY 2563 GATCTACAACTCTCAAAATGCAATTTTGTATGATACCTAGTGTGTTGTTCAAGTGTGCTT 2622
DB 664 AASRYVANAANAKWYTYKBAANNAYTHANNWGCWNNATDTRTWKNNNNNAGTQKN 723
QY 2623 GCGAAGAAAAGTTAGATAATTTTGAAGCTATGGGACCAAAATTAATTTATGATGATG 2682
DB 724 NNNNAKNSAAAKVYAAAANKAAKHWKANKWAMRGHADAAABTTDKENNGAYTKYTT 783
QY 2683 CTCGAACCTAGTTGATGATGATGAGCCCAATAATAGACACTAGTTCATGCGGTGTTCT 2742
DB 784 NNNNTYRGVVTNTAARDGWANNNNNNNNNNNNNNNNNNNNWVWVWYANYGTNNNNNNNN 843

```

```

QY 2743 TGTATAGTACTAGCTAGTATAAATCTTTTCAAGTTGTAGCTACTACTTTAGCTTACTCTCC 2802
DB 844 NAYAWTNKWYTTTDDRWRBAYTNNNNNNRMYIYAYYMSDTCDAWMKWDATKM 903
QY 2803 GCATATTACAATCAATAGAAATCGGAAGTACTATAAACGGAGCCTATA 2852
DB 904 NNATTYNRGTAWRTNNNNNNMTKTKYBHAANNNNNNNGKGTACTAHTWVW 953

RESULT 11
US-09-806-708B-22/c
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAEI promoters
US-09-806-708B-22

Query Match
Best Local Similarity 1.5%; Score 46; DB 4; Length 1141;
Matches 115; Conservative 352; Mismatches 546; Indels 2; Gaps 2;

QY 678 TATCTCTAAAGGGGAACGAATGGATGGTGGGACACGTGGGGAGACACCGAAGGACATG 737
DB 1093 TAWTNHAKRATWCWYVYWTGTRRCWRYAMRTWYTRSNANWSCATKBMWTKWYA 1034
QY 738 CCGAGGAGGACACAAAGCTTCAGCAGGCTCCAGACTCTCAGAGAAGAAAGCTCA 797
DB 1033 TKYRTAWYMWCAWRNNNNWCATNGYAKSCATNNAMWYATTWAAAYAAAKWARWAGNNNR 974
QY 798 CGGACAGGTTGGGCTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 857
DB 973 MYGAAAGKWKGCMAAMATMGBWADTAGKMCNNNNNNNTDVRMMAMKAKNNNNNNAYWT 914
QY 858 GCTGAAATCGACCGCGCGGACCAACAGAGGTCAGCTCGGCACCTCCGCTCTCCGAGC 917
DB 913 ACYNRAATNNKATMHWKTHGASHKRTTRHRTCRRTKYNNNNNNNARTVYVYHHAAR 854
QY 918 GCATGAGTGACCGCTTCGTCGCGGCTTCCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 977
DB 853 WMAWATRTNNNNNNNNNACNRTTWABWKHSKNNNNNNNNNNNNNNNNNNNTWCHYTANA 794
QY 978 TCACGGGACCGCTGAAACCAATCAGACGTTCCCTTTACAGGGGAAAGGCAAGCTCTGA 1037
DB 793 BBYRANNNNAARMAARTCNMYMHAAVTTTHDWCYKTMNTWYMDMTTMBTTTTRNMT 734
QY 1038 TAACTCTCTGTTTCCATCGCTCTTAACCGGAAGAGCGGACGCAACAGACTTAGAGTC 1097
DB 733 TSTNNNNNNNNMAC-TNNNNNNMWKAYAHATNNWGCWNNNTDARTNNNTTVVBRWM 675
QY 1098 TATTGTTTCGAAATTTTACTCTCAAAAAGCTAGCTTTTATAGCGGGGATAAAAGCT 1157
DB 674 INTKRWYSTTRRHHTTGTATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 615
QY 1158 ATCATGTCGCGGACGCTTTTAAATTAATTAACCTTATACCATATGATATATCATGCTCA 1217
DB 614 KVKWRDITCTTVDVWADSWVWVWYANWRCRDVYTRNNYCKSYAHSYVWYSSNNAMWYRR 555

```


US-09-949-016-150744
; Sequence 150744, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 150744
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-150744

Query Match 1.4%; Score 40.8; DB 4; Length 601;
Best Local Similarity 48.3%; Pred. No. 0.088;
Matches 114; Conservative 0; Mismatches 122; Indels 0; Gaps 0;
QY 1198 ATGAATATCATGTCGAACCTATGAGGATGATACCTTTCTGAACTGATTCGCGTGAGTTATT 1257
Db 50 ATGTTTAGGTTTTTAAGAAATGCGCAAGTTGTTTCCAGAAATTTTCCCTAGAAATATTC 109
QY 1258 AAATTGTACTTTAGTTGTTGAGCATGAAGCTCTGAACATGAAATTTATGATGATTTGT 1317
Db 110 AGAAGGTACATTTTACATTCCTATCAAAATATATGAATGAGGACTTCTCTGCATTCC 169
QY 1318 GCGTTGTAGCTACCTCCGCTTACATTTAGTTGGTATCATAAATATATATATATATCAT 1377
Db 170 AGCACTTGATTTGCTACCACTTTTCAATTTAGCCACTGTGATAGATATGATCTCATTTGTC 229
QY 1378 ATAAATTTGATCAACTTGAGATGCTTTGACTCTTCAAGATTTCTTGAATGACTTAT 1433
Db 230 TTAAATTTGCATTTCCCTTAATGGCTGATAGTGTAAACATCATTTTGTCTGTGCTTAT 285

RESULT 15
US-09-949-016-15974/c
; Sequence 15974, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15974
; LENGTH: 294836
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(294836)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15974

Query Match 1.4%; Score 40.8; DB 4; Length 294836;
Best Local Similarity 48.3%; Pred. No. 10; Mismatches 0; Gaps 0;
Matches 114; Conservative 0; Mismatches 122; Indels 0; Gaps 0;
QY 1198 ATGAATATCATGTCGAACCTATGAGGATGATACCTTTCTGAACTGATTCGCGTGAGTTATT 1257
Db 131961 ATGTTTAGGTTTTTAAGAAATGCGCAAGTTGTTTCCAGAAATTTTCCCTAGAAATATTC 131902
QY 1258 AAATTGTACTTTAGTTGTTGAGCATGAAGCTCTGAACATGAAATTTATGATGATTTGT 1317
Db 131901 AGAAGGTACATTTTACATTCCTATCAAAATATATGAATGAGGACTTCTCTGCATTCC 131842
QY 1318 GCGTTGTAGCTACCTCCGCTTACATTTAGTTGGTATCATAAATATATATATATATCAT 1377
Db 131841 AGCACTTGATTTGCTACCACTTTTCAATTTAGCCACTGTGATAGATATGATCTCATTTGTC 131782
QY 1378 ATAAATTTGATCAACTTGAGATGCTTTGACTCTTCAAGATTTCTTGAATGACTTAT 1433
Db 131781 TTAAATTTGCATTTCCCTTAATGGCTGATAGTGTAAACATCATTTTGTCTGTGCTTAT 131726

Search completed: February 16, 2005, 10:31:38
Job time : 494 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 16, 2005, 03:45:17 ; Search time 1981 Seconds
(without alignments)
8983.234 Million cell updates/sec

Title: US-10-751-612-1

Perfect score: 3016

Sequence: 1 tctagacataggcattgta.....gggctatgtcaagtcctatgg 3016

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5378673 seqs, 2950229984 residues

Total number of hits satisfying chosen parameters: 10757346

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	3016	100.0	3016	18	US-10-751-612-1		Sequence 1, Appli
2	120.8	4.0	576	18	US-10-425-115-151068		Sequence 151068, A
3	120.2	4.0	622	18	US-10-767-701-19690		Sequence 19690, A
4	118.6	3.9	3688	9	US-09-866-153-10		Sequence 10, Appli
5	118.6	3.9	3691	9	US-09-866-153-3		Sequence 3, Appli
6	118.6	3.9	3174	9	US-09-866-153-8		Sequence 8, Appli
7	100.8	3.3	1346	18	US-10-767-701-11451		Sequence 11451, A
8	96.4	3.2	398	18	US-10-767-701-18110		Sequence 18110, A
9	89.4	3.0	1318	18	US-10-767-701-10015		Sequence 10015, A
10	89.4	3.0	2032	18	US-10-425-115-42192		Sequence 42192, A
11	88	2.9	1894	17	US-10-425-114-32756		Sequence 32756, A

c	12	1894	2.9	1894	18	US-10-425-115-183742	Sequence 183742, A
	13	1285	2.9	1285	17	US-10-425-114-24208	Sequence 24208, A
	14	1360	2.9	1360	18	US-10-425-115-142566	Sequence 142566, A
	15	1382	2.9	1382	18	US-10-425-115-165368	Sequence 165368, A
	16	690	2.8	690	18	US-10-767-701-10788	Sequence 10788, A
	17	78.8	2.6	586	18	US-10-767-701-17017	Sequence 17017, A
	18	66.8	2.2	2000	17	US-10-260-238-1707	Sequence 1707, A
	19	66	2.2	129	18	US-10-425-115-133113	Sequence 133113, A
	20	65.4	2.2	1334	17	US-10-333-006-7	Sequence 7, Appli
	21	62.6	2.1	1341	17	US-10-425-114-27759	Sequence 27759, A
	22	62.2	2.1	1663	18	US-10-425-115-25020	Sequence 25020, A
	23	62	2.1	701	17	US-10-389-566-4	Sequence 4, Appli
	24	62	2.1	1626	18	US-10-425-115-25018	Sequence 25018, A
	25	61.4	2.0	439	18	US-10-437-963-31959	Sequence 31959, A
	26	61.2	2.0	290	9	US-09-923-876-3721	Sequence 3721, A
	27	61.2	2.0	290	10	US-09-923-876-3721	Sequence 3721, Ap
	28	61.2	2.0	1374	17	US-10-425-114-3568	Sequence 3568, Ap
	29	61.2	2.0	1705	18	US-10-425-115-124251	Sequence 124251, A
	30	60.8	2.0	439	18	US-10-437-963-42140	Sequence 42140, A
	31	60.8	2.0	635	18	US-10-767-701-517	Sequence 517, App
	32	60.8	2.0	1475	18	US-10-437-963-53393	Sequence 53393, A
	33	60.6	2.0	1894	17	US-10-425-114-32756	Sequence 32756, A
	34	60.6	2.0	1894	18	US-10-425-115-183742	Sequence 183742, A
	35	59.8	2.0	1309	17	US-10-425-114-35130	Sequence 35130, A
	36	59.6	2.0	631	18	US-10-767-701-24985	Sequence 24985, A
	37	59.2	2.0	255	18	US-10-437-963-93872	Sequence 93872, A
	38	58	1.9	326	18	US-10-437-963-66554	Sequence 66554, A
	39	57.8	1.9	442	18	US-10-437-963-95390	Sequence 95390, A
	40	57.6	1.9	305	18	US-10-437-963-17495	Sequence 17495, A
	41	57.6	1.9	464	18	US-10-437-963-2371	Sequence 2371, Ap
	42	57.6	1.9	2000	17	US-10-260-238-1707	Sequence 1707, A
	43	57.4	1.9	442	18	US-10-437-963-42065	Sequence 42065, A
	44	57.2	1.9	441	18	US-10-437-963-31961	Sequence 31961, A
	45	57	1.9	452	18	US-10-437-963-30623	Sequence 30623, A

ALIGNMENTS

RESULT 1

US-10-751-612-1
; Sequence 1, Application US/10751612
; Publication No. US2005005322A1
; GENERAL INFORMATION:
; APPLICANT: Mirkov, T. Erik
; APPLICANT: Damaj, Mona B.,
; APPLICANT: Reddy, Avutu,
; APPLICANT: Thomas, Terry L.,
; APPLICANT: Rathore, Keerti S.,
; APPLICANT: Emami, Chandrakanth,
; APPLICANT: Kumpatla, Siva Prasad
; TITLE OF INVENTION: STEM-REGULATED, PLANT DEFENSE PROMOTER
; TITLE OF INVENTION: AND USE THEREOF IN TISSUE-SPECIFIC EXPRESSION IN MONOCOTS
; FILE REFERENCE: 017575.0774
; CURRENT APPLICATION NUMBER: US/10/751.612
; CURRENT FILING DATE: 2004-01-05
; PRIOR APPLICATION NUMBER: 60/437,890
; PRIOR FILING DATE: 2003-01-03
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3016
; TYPE: DNA
; ORGANISM: Sugarcane
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)...(3012)
; OTHER INFORMATION: o-methyltransferase promoter
; FEATURE:
; NAME/KEY: CAAT_signal
; LOCATION: (2661)...(2664)
; FEATURE:
; NAME/KEY: TATA_signal

LOCATION: (2849).... (2855)									
US-10-751-612-1									
Query Match 100.0%; Score 3016; DB 18; Length 3016;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 3016; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	TC	TAGAGCATAGGCAATGTA	AAAGCGGTATGCCTCTT	CTCAGTGAGAAATTT	CATACCA	60		
DB	1	TC	TAGAGCATAGGCAATGTA	AAAGCGGTATGCCTCTT	CTCAGTGAGAAATTT	CATACCA	60		
QY	61	AC	CTTAGGTATCCTGTCCTT	CCATAGAAATTTCT	TACCTGAGTAGTT	CGGTCTGTTGGAT	120		
DB	61	AC	CTTAGGTATCCTGTCCTT	CCATAGAAATTTCT	TACCTGAGTAGTT	CGGTCTGTTGGAT	120		
QY	121	TT	GTAGCGGTTTCATGCA	AAATPAAGTTAGA	ATCGTCAAACTTG	CAATGGAGTTAA	180		
DB	121	TT	GTAGCGGTTTCATGCA	AAATPAAGTTAGA	ATCGTCAAACTTG	CAATGGAGTTAA	180		
QY	181	TT	TGAATATATTTGCAT	AGCAAAACAAATAT	TAGATTGAT	TGTAATGCAATATGAC	240		
DB	181	TT	TGAATATATTTGCAT	AGCAAAACAAATAT	TAGATTGAT	TGTAATGCAATATGAC	240		
QY	241	TT	GCAATTTCTAACTCT	ATTGCTACTGTGCC	AGATGAAGAATGTT	GATCTGGAGAAGTTT	300		
DB	241	TT	GCAATTTCTAACTCT	ATTGCTACTGTGCC	AGATGAAGAATGTT	GATCTGGAGAAGTTT	300		
QY	301	TG	TGAGAAATGTGACAA	CAACGGAGTCAAT	CAAGATTTCTGG	TACCCGCGAGAAATCG	360		
DB	301	TG	TGAGAAATGTGACAA	CAACGGAGTCAAT	CAAGATTTCTGG	TACCCGCGAGAAATCG	360		
QY	361	GC	CTCATCTAGTTAGC	CTCTGAGCATGGGG	GAATGGCTGAGATG	CCCCCATGTAG	420		
DB	361	GC	CTCATCTAGTTAGC	CTCTGAGCATGGGG	GAATGGCTGAGATG	CCCCCATGTAG	420		
QY	421	TC	GTAGGCATCGAGAG	TACTGGCTGAGATG	CCATTTGTTGAT	CGAGAGAAACGAG	480		
DB	421	TC	GTAGGCATCGAGAG	TACTGGCTGAGATG	CCATTTGTTGAT	CGAGAGAAACGAG	480		
QY	481	AGA	ATGCTAGTCTAATA	TACCTTCGTATGCT	TAAACCACTAT	TATAATTTGGACCAT	540		
DB	481	AGA	ATGCTAGTCTAATA	TACCTTCGTATGCT	TAAACCACTAT	TATAATTTGGACCAT	540		
QY	541	TTTT	CACATGCTAGCGC	CTTTGCTCTTATTT	TAATTAATTTCAAT	TGGTCCGATAGCATG	600		
DB	541	TTTT	CACATGCTAGCGC	CTTTGCTCTTATTT	TAATTAATTTCAAT	TGGTCCGATAGCATG	600		
QY	601	TG	AACTGGGAGACGGT	TCCGTGCGACGGCT	CCGTTTCTTTGAT	CGGTACGGCGTGAGC	660		
DB	601	TG	AACTGGGAGACGGT	TCCGTGCGACGGCT	CCGTTTCTTTGAT	CGGTACGGCGTGAGC	660		
QY	661	GAG	AAAGTGAGGCGCT	ATCTTAAAGGGNAC	CAATGGATGGT	GACAGTGGGGA	720		
DB	661	GAG	AAAGTGAGGCGCT	ATCTTAAAGGGNAC	CAATGGATGGT	GACAGTGGGGA	720		
QY	721	GA	CCGAGGACATGCC	GAGGACACAAGCTT	CAGAGGGCTCTC	CAGACTCTCA	780		
DB	721	GA	CCGAGGACATGCC	GAGGACACAAGCTT	CAGAGGGCTCTC	CAGACTCTCA	780		
QY	781	GA	GAAGAAGAGCTC	AGCGACGGTGGG	TGTTCTTGCTGT	CTCGTGTG	840		
DB	781	GA	GAAGAAGAGAGCT	CACGGCA	CGGTTGCGGTGGT	CTTGCTGTCTCGTGTG	840		
QY	841	CA	CGTTTCTGTGAT	CACTGAAATCG	ACCGCGGCGGAC	CAACAGAGGTCA	900		
DB	841	CA	CGTTTCTGTGAT	CACTGAAATCG	ACCGCGGCGGAC	CAACAGAGGTCA	900		
QY	901	CA	CTCCGTCTCGAG	CCATGAGTGA	CCGTTCTCGCGGGT	CTCTTCTCGTGTG	960		
DB	901	CA	CTCCGTCTCGAG	CCATGAGTGA	CCGTTCTCGCGGGT	CTCTTCTCGTGTG	960		
QY	961	CG	TGACCCCTCTGGT	TTCACCGGAC	CCCTGAACCAAT	CAGACGTTCCCTTTACAGG	1020		
DB	961	CG	TGACCCCTCTGGT	TTCACCGGAC	CCCTGAACCAAT	CAGACGTTCCCTTTACAGG	1020		
Db	961	CG	TGCAACGCTCTGCGT	TTTCAACGGCAC	CCCTGAACCAAT	CAGACGTTCCCTTTACAGG	1020		
QY	1021	GAA	GGGCAAGTCTGATA	AACCTCTCTCTT	CCATCGTCTT	CAACCGGAGCGGAC	1080		
DB	1021	GAA	GGGCAAGTCTGATA	AACCTCTCTCTT	CCATCGTCTT	CAACCGGAGCGGAC	1080		
QY	1081	GC	CAAGACTT	PAGAGTCTATT	TGTTGCGAAATTTT	TACTCTCACA	1140		
DB	1081	GC	CAAGACTT	PAGAGTCTATT	TGTTGCGAAATTTT	TACTCTCACA	1140		
QY	1141	AG	CGGGCATAAA	AGCTATCAT	GTGACCGGCA	CGTTTAAATTT	1200		
DB	1141	AG	CGGGCATAAA	AGCTATCAT	GTGACCGGCA	CGTTTAAATTT	1200		
QY	1201	AA	TATCATGT	CGAACTAT	GAAGTAT	CTTCTGAACG	1260		
DB	1201	AA	TATCATGT	CGAACTAT	GAAGTAT	CTTCTGAACG	1260		
QY	1261	TT	GTACTTTT	TAGTTGTT	GAGCATGAAG	GTCTGAATTT	1320		
DB	1261	TT	GTACTTTT	TAGTTGTT	GAGCATGAAG	GTCTGAATTT	1320		
QY	1321	TT	GTAGACTACT	CCGCTCTAC	ATTAGTTAGT	TGTTATATTT	1380		
DB	1321	TT	GTAGACTACT	CCGCTCTAC	ATTAGTTAGT	TGTTATATTT	1380		
QY	1381	AA	TTTGAATCAA	CTTTGAGATG	CTTTGAC	CTTTTGAAGAT	1440		
DB	1381	AA	TTTGAATCAA	CTTTGAGATG	CTTTGAC	CTTTTGAAGAT	1440		
QY	1441	GG	TAGGAGTAGT	TTTCTAAG	GGCAGTCTCAG	TGGGTTTCAT	1500		
DB	1441	GG	TAGGAGTAGT	TTTCTAAG	GGCAGTCTCAG	TGGGTTTCAT	1500		
QY	1501	TA	ATAAGCTGAT	GTACACCG	TATGATGAAG	GAGAGATGAT	1560		
DB	1501	TA	ATAAGCTGAT	GTACACCG	TATGATGAAG	GAGAGATGAT	1560		
QY	1561	TAG	AGAGTTTCA	TGGGATGA	AATCTTCTTCA	CTTTTCCAAA	1620		
DB	1561	TAG	AGAGTTTCA	TGGGATGA	AATCTTCTTCA	CTTTTCCAAA	1620		
QY	1621	GT	AGAGGCGCAT	GAAATCT	TAGTGAC	CTGACCTAAGAT	1680		
DB	1621	GT	AGAGGCGCAT	GAAATCT	TAGTGAC	CTGACCTAAGAT	1680		
QY	1681	TG	TTTCAAAATCT	GCATGCAT	GCATGCTTTG	AACTTAACT	1740		
DB	1681	TG	TTTCAAAATCT	GCATGCAT	GCATGCTTTG	AACTTAACT	1740		
QY	1741	AC	CATGCATGCA	AAACGGGCGGT	GCACGCAAAAG	ATTTGATGAG	1800		
DB	1741	AC	CATGCATGCA	AAACGGGCGGT	GCACGCAAAAG	ATTTGATGAG	1800		
QY	1801	TA	GTAAAATGCT	TTTGGCTT	CATCACCGGCTT	AAATGCTCGAC	1860		
DB	1801	TA	GTAAAATGCT	TTTGGCTT	CATCACCGGCTT	AAATGCTCGAC	1860		
QY	1861	AG	TCAAGGTTGT	CGCTTAA	CAAACTGGGGT	TCAATGT	1920		
DB	1861	AG	TCAAGGTTGT	CGCTTAA	CAAACTGGGGT	TCAATGT	1920		
QY	1921	CG	CGCTGGAGG	ATTAGATCA	CACTTCAAT	TATCTTAGGG	1980		
DB	1921	CG	CGCTGGAGG	ATTAGATCA	CACTTCAAT	TATCTTAGGG	1980		
QY	1981	CT	CTAAACTAG	TTTTAT	TGTACCGT	TGAGGAGG	2040		
DB	1981	CT	CTAAACTAG	TTTTAT	TGTACCGT	TGAGGAGG	2040		
QY	2041	CT	AACTGAG	AGAGCTAT	TTTTTTTT	TGCTCCCAAT	2100		
DB	2041	CT	AACTGAG	AGAGCTAT	TTTTTTTT	TGCTCCCAAT	2100		

Db 364 CTCTGGGTACATGAGGAGAACTGCTCTTCCATGTGCTGTTGCTCTCGTCAGGCATGGGGCAA 423
QY 398 TTGGCTGAGATGCCCCCATGTA-----GTCTCAGGCATGGAGA 436
Db 424 TTGGCTGAGATGCCCCCGTATAGCTGGTCTTGTGTGGCACTGGTGTCTGTCAGGCATGGAGA 483
QY 437 GTACTGGCTGAGATGCCATCTGTTGTGTAGATCGAGAGAAAACGAGAGAAGATGCTAGTCTAA 496
Db 484 GAATTTGGGTGAGATTGAGAGAAACGAGAGAAGATGCTAATCTAATTACCTTCCTCGTAT 543
QY 497 TAATACCCCTCCGTATGCTAACCACTATTAATAATGGCACCATTATTTACATGCTA 553
Db 544 CTAGATGATAAATTGATGAACCAACTATTATAATTAGCGCCATTATTTACAGGCAA 600

RESULT 4

US-09-866-153-10/c
; Sequence 10, Application US/09866153
; Patent No. US2002004615A1
; GENERAL INFORMATION:
; APPLICANT: Albert, Henrik H.
; APPLICANT: Wei, Hairong
; TITLE OF INVENTION: PLANT PROMOTER SEQUENCES AND METHODS OF USE THEREOF
; FILE REFERENCE: UH-03648
; CURRENT APPLICATION NUMBER: US/09/866,153
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/270,976
; PRIOR FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 3688
; TYPE: DNA
; ORGANISM: Saccharum Hybrid Cultivar 32-8560
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9)
; OTHER INFORMATION: The "n" at position 9 is any nucleotide.
US-09-866-153-10

Query Match 3.9%; Score 118.6; DB 9; Length 3688;
Best Local Similarity 86.0%; Pred. No. 7.1e-21;
Matches 166; Conservative 0; Mismatches 24; Indels 3; Gaps 3;
QY 1451 AGTTTCTTAAGCCAGTCTCAGTGGGGTTTCATCAGAGTTTCATGACATTAATAAGCT 1510
Db 1914 AGTTTCTGAGCCCGTCTCAGT-GGATTTTCATCAGAGTTTCATGACATTAATAAGCT 1856
QY 1511 GATGTGACACCGTATTGATGAAGAGAGAGATGATAAGAGTTTCATGCGAGTAGAGAGAGT 1570
Db 1855 GATGTGGCACCGTATTGATGAAGAGAGAGATGATAAGAGTTTCATGCGAATTAGAGAGAGT 1796
QY 1571 TTTCATGGGATGAAACTCTT-CTTCACGTGTTTCCAAAATAT-AGATGCAATTGGTAAGAGG 1628
Db 1795 TTTCAGAAATGAAACTCTCTCTGCACTGTTTCCAAAATATGGGTGCAATTAATAACATG 1736
QY 1629 GCCATGAATCTC 1641
Db 1735 GCCATAAAATCCC 1723

RESULT 5

US-09-866-153-3/c
; Sequence 3, Application US/09866153
; Patent No. US2002004615A1
; GENERAL INFORMATION:
; APPLICANT: Albert, Henrik H.
; APPLICANT: Wei, Hairong
; TITLE OF INVENTION: PLANT PROMOTER SEQUENCES AND METHODS OF USE THEREOF
; FILE REFERENCE: UH-03648
; CURRENT APPLICATION NUMBER: US/09/866,153
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/270,976

; PRIOR FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 3691
; TYPE: DNA
; ORGANISM: Saccharum Hybrid Cultivar H32-8560
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9)
; OTHER INFORMATION: The "n" at position 9 is any nucleotide.
; NAME/KEY: misc_feature
; LOCATION: (3613)
; OTHER INFORMATION: The "n" at position 3613 is any nucleotide.
US-09-866-153-3

Query Match 3.9%; Score 118.6; DB 9; Length 3691;
Best Local Similarity 86.0%; Pred. No. 7.1e-21;
Matches 166; Conservative 0; Mismatches 24; Indels 3; Gaps 3;
QY 1451 AGTTTCTTAAGCCAGTCTCAGTGGGGTTTCATCAGAGTTTCATGACATTAATAAGCT 1510
Db 1914 AGTTTCTGAGCCCGTCTCAGT-GGATTTTCATCAGAGTTTCATGACATTAATAAGCT 1856
QY 1511 GATGTGACACCGTATTGATGAAGAGAGAGATGATAAGAGTTTCATGCGAGTAGAGAGAGT 1570
Db 1855 GATGTGGCACCGTATTGATGAAGAGAGAGATGATAAGAGTTTCATGGAATTAGAGAGAGT 1796
QY 1571 TTTCATGGGATGAAACTCTT-CTTCACGTGTTTCCAAAATAT-AGATGCAATTGGTAAGAGG 1628
Db 1795 TTTCAGAAATGAAACTCTCTCTGCACTGTTTCCAAAATATGGGTGCAATTAATAACATG 1736
QY 1629 GCCATGAATCTC 1641
Db 1735 GCCATAAAATCCC 1723

RESULT 6

US-09-866-153-8/c
; Sequence 8, Application US/09866153
; Patent No. US2002004615A1
; GENERAL INFORMATION:
; APPLICANT: Albert, Henrik H.
; APPLICANT: Wei, Hairong
; TITLE OF INVENTION: PLANT PROMOTER SEQUENCES AND METHODS OF USE THEREOF
; FILE REFERENCE: UH-03648
; CURRENT APPLICATION NUMBER: US/09/866,153
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/270,976
; PRIOR FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 5174
; TYPE: DNA
; ORGANISM: Saccharum Hybrid Cultivar H32-8560
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9)
; OTHER INFORMATION: The "n" at position 9 is any nucleotide.
; NAME/KEY: misc_feature
; LOCATION: (4318)
; OTHER INFORMATION: The "n" at position 4318 is any nucleotide.
; NAME/KEY: misc_feature
; LOCATION: (4546)
; OTHER INFORMATION: The "n" at position 4546 is any nucleotide.
; NAME/KEY: misc_feature
; LOCATION: (4890)
; OTHER INFORMATION: The "n" at position 4890 is any nucleotide.
; NAME/KEY: misc_feature
; LOCATION: (5117)
; OTHER INFORMATION: The "n" at position 5117 is any nucleotide.
; NAME/KEY: misc_feature

Db 1308 CATCAAAAC 1318

RESULT 10

US-10-425-115-42192
; Sequence 42192, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 42192
; LENGTH: 2032
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_138478C.1
US-10-425-115-42192

Query Match 3.0%; Score 89.4; DB 18; Length 2032;

Best Local Similarity 68.9%; Pred. No. 5e-13;
Matches 151; Conservative 0; Mismatches 66; Indels 2; Gaps 2;

QY 1457 CTAAGGCCAGTCTCAGTGG-GGTTTCATCAGAGTTCATGGACATTAAATAAGCTGATGT 1515

DB 498 CTAAGACTAGTCTCTGTGGTTCACGAGGATTATGGCATTAATATGTTGATGT 557

QY 1516 GACACCGTATTGATGAAGAGAGAGATGATAAGAGTTTTCATGCGAGPAGAGAGTTTCAT 1575

DB 558 GGCATATTATTAAACGAAGAGAGATAAGGTAAAGTTTTTATCGAATGAATGAGTTCCAC 617

QY 1576 GGG-GATGAACCTCTTCCACTGTTCCAAATATAGATGATGTTGTAAGGGGCCATG 1634

DB 618 GGGCGATGAACCTTATGTGCACTGCTTCAACATATCGAGTCTTGGGAACATTGACATA 677

QY 1635 AAATCTCTAGTACACTGACCTTAAGATGAGATTGACTCT 1673

DB 678 AAATCCCCTGAACCTGGCTTAATAAATAGCTAATCT 716

RESULT 11

US-10-425-114-32756/c
; Sequence 32756, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 32756
; LENGTH: 1894
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-2MFLM017052G10_FLI
US-10-425-114-32756

Query Match

2.9%; Score 88; DB 17; Length 1894;

Best Local Similarity 83.3%; Pred. No. 1.2e-12;
Matches 100; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 2452 AAAGAGTTAAATGATGATGAGTCTTGAATCTTCTGAGAGTGCCACTTAGGTCCCAA 2511

DB 666 AAAGAGTTAAATGATGAGTCTTGAACCTTCTGAGAGTGCCACTTAGATACATGA 607

QY 2512 ACTCTCAAATTCATTTTTCACACCCCTAAATGTTTCAAGTGTGCCACTTAGATCTACAA 2571

DB 606 ACTCTCAATGATATTTTGGCACCCCTAAATATGTTTAAAGTGTGTCACTTAGGTCCATAA 547

RESULT 12

US-10-425-115-183742/c
; Sequence 183742, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 183742
; LENGTH: 1894
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_99156C.1
US-10-425-115-183742

Query Match 2.9%; Score 88; DB 18; Length 1894;

Best Local Similarity 83.3%; Pred. No. 1.2e-12;
Matches 100; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 2452 AAAGAGTTAAATGATGATGAGTCTTGAATCTTCTGAGAGTGCCACTTAGGTCCCAA 2511

DB 666 AAAGAGTTAAATGATGAGTCTTGAACCTTCTGAGAGTGCCACTTAGATACATGA 607

QY 2512 ACTCTCAAATTCATTTTTCACACCCCTAAATGTTTCAAGTGTGCCACTTAGATCTACAA 2571

DB 606 ACTCTCAATGATATTTTGGCACCCCTAAATATGTTTAAAGTGTGTCACTTAGGTCCATAA 547

RESULT 13

US-10-425-114-24208
; Sequence 24208, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 24208
; LENGTH: 1285
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3608-006-C5_FLI
US-10-425-114-24208


```

; LENGTH: 1382
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_82401C.1
US-10-425-115-165368

Query Match          2.9%; Score 86.2; DB 18; Length 1382;
Best Local Similarity 87.9%; Pred. No. 2.9e-12;
Matches 94; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy      2910  CGCCATGGCGCTCAGCAAGGAGCAACACACAGCACTGATCAGCAGCGCCTGCTGGATGC 2969
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      14    CGCCATGGCGCTCAGCAAGGAGCAAGCGCAGCACTGATCAGCAGCGCTTGTCTGGATGC 73

Qy      2970  TCAGCTCCAGCTCTGGCACCACACACCCCTGGGCTATGTCAAGTCCATGG 3016
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      74    TCAGCTGCAGCTCTGGCACCACACACCATCGCTTTTGTCAAGTCCATGG 120

Search completed: February 16, 2005, 11:04:46
Job time : 1985 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2005, 22:03:37 ; Search time 9452 Seconds
(without alignments)
12145.772 Million cell updates/sec

Title: US-10-751-612-1

Perfect score: 3016

Sequence: 1 tctagagcataggcatgtga.....gggctagtccaagtcctatgg 3016

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_hcc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gse1:*
9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	320.4	10.6	322	6	CA113147	CA113147 SCEZLB100
2	275.2	9.1	583	6	CA118968	CA118968 SCBGLR111
3	254.6	8.4	556	6	CA248733	CA248733 SCBFL110
4	164.6	5.5	674	6	CA203163	CA203163 SCRFLL102
5	161.6	5.4	563	6	CA128031	CA128031 SCAGLR201
6	158	5.2	664	6	CA154373	CA154373 SCVPRZ204
7	157.8	5.2	594	6	CA126310	CA126310 SCBGLR104
8	154.4	5.1	573	6	CA118511	CA118511 SCBGLR109
9	152.2	5.0	634	9	CW512557	CW512557 115_1_105
10	149.2	4.9	400	6	CA186194	CA186194 SCBGLR111
11	148.6	4.9	717	9	CL192172	CL192172 104_414_1
12	147.8	4.9	669	6	CA220522	CA220522 SCRFL402
13	147.4	4.9	740	9	CL191874	CL191874 104_414_1
14	141.8	4.7	1273	8	BZ693362	BZ693362 SP_Ba003
15	139.4	4.6	787	8	BZ342445	BZ342445 1c83e09.g
16	136.2	4.5	732	8	BZ715359	BZ715359 PUFAD81TD
17	136.2	4.5	907	9	CG369374	CG369374 CG0GR27TV
18	136.2	4.5	921	9	CG068037	CG068037 PUIC82TB
19	133.8	4.4	591	8	BZ692106	BZ692106 SP_Ba001
20	130.6	4.3	664	8	BZ692106	BZ692106 SP_Ba001
21	129	4.3	680	8	BZ626898	BZ626898 1h49a06.b
22	128.8	4.3	623	9	CL193585	CL193585 104_417_1
23	126.6	4.2	631	9	CW512200	CW512200 115_1_105
24	126.4	4.2	708	8	BZ343488	BZ343488 ho55h06.b

25	126.2	4.2	734	6	CF070857	CF070857 FEL_14_B0
26	125.2	4.2	622	8	BZ629605	BZ629605 ih67c10.b
27	123.6	4.1	624	6	CA186182	CA186182 SCGST311
28	122.8	4.1	559	6	CA118575	CA118575 SCBGLR110
29	120.8	4.0	419	9	CL1701972	CL1701972 SP_Ba007
30	120.8	4.0	747	9	CL152042	CL152042 104_335_1
31	119.6	4.0	1246	8	BZ694416	BZ694416 SP_Ba004
32	118.6	3.9	925	9	CG295925	CG295925 OG0FA16TV
33	115.6	3.8	334	8	BZ629606	BZ629606 ih67c10.g
34	115.6	3.8	389	8	BZ989929	BZ989929 PUDFG82TD
35	114.2	3.8	512	5	BQ536419	BQ536419 STEM4_A
36	114.2	3.8	513	5	BQ536474	BQ536474 STEM4_1_D
37	114.2	3.8	625	5	BQ537299	BQ537299 STEM2_23
38	114.2	3.8	637	5	BQ536451	BQ536451 STEM4_1_B
39	113.8	3.8	770	9	CL167477	CL167477 104_364_1
40	113.8	3.8	782	9	CL167479	CL167479 104_364_1
41	113.4	3.7	726	8	CL154621	CL154621 104_340_1
42	113	3.7	630	8	CC059054	CC059054 1118f07.b
43	111.2	3.7	637	6	CA195175	CA195175 SCEZSB109
44	111.2	3.7	660	5	BQ536496	BQ536496 STEM4_1_E
45	110.8	3.7	658	6	CA210698	CA210698 SCEPSB113

ALIGNMENTS

RESULT 1
CA113147
LOCUS SCEZLB1009D09.g LB1 Saccharum officinarum cdna clone SCEZLB1009D09
DEFINITION 322 bp mRNA linear EST 23-SEP-2003
5', mRNA sequence.
ACCESSION CA113147
VERSION CA113147.1 GI:34966454
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
complex.
REFERENCE 1 (bases 1 to 322)
AUTHORS Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.
TITLE The libraries that made SUCEST
JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 009 row: D column: 09
Seq primer: T7 Promoter Primer.
Location/Qualifiers
1. 322
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCEZLB1009D09"
/lab_host="DH10B"
/clone_lib="LB1"

note="Organ: Lateral buds from field grown adult plants;
Vector: pSPort1; Site 1: SalI; Site 2: NotI; An
unidirectional cDNA library generated from [lateral buds
from field grown adult plants]. cDNA was prepared from
polyA+ mRNA using SuperScript Plasmid System kit
(Invitrogen). The double-strand cDNAs were fractionated
in a sepharose CL-2B 40cm-columns and fragments sizing
between 0.8 and 1.5 kb were directionally cloned into the
vector. Details of each source of RNA and library

```

FEATURES
source
seq_frames
seq_frames.qualifiers
location/Qualifiers
1..556
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCSBF1104B06"
/lab_host="DH10B"

```

/clone_lib="FL1"
/note="Organ: Inflorescence at beginning of development (lcm-long); Vector: pSport1; Site_1: SalI; Site_2: NotI; An unidirectional cDNA library generated from [Inflorescence at beginning of development (lcm-long)]. cDNA was prepared from polyA+ mRNA using SuperScript plasmid system kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://sucust.lad.ic.unicamp.br/public>"

ORIGIN

Query Match 8.4%; Score 254.6; DB 6; Length 556;
Best Local Similarity 97.5%; Pred. No. 5.5e-58;
Matches 270; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

QY 272 AGATGAAGAATGTCATCTGGAGAAGTTTGTGAGAATGTGACAACACGGGAGGTGATA 331
Db |||||
QY 277 AGATGAAGAATGTCATCTGGAGAAGTTTGTGAGAATGTGACAACACAGGAGGTGATA 218
Db |||||
QY 332 TCAAGATTCTGGGTACCGCGGAGAAATCGGCTCCATGTAGTTAGCTTCGTCAGGCATGG 391
Db |||||
QY 217 TCAAGATTCTGGGTACCGCGGAGAAATCGGCTCCATGTAGTTAGCTTCGTCAGGCATGG 158
Db |||||
QY 392 GGGGAATTGGCTGAGATGCCCATGTAGTGTCTCAGGCATGGAGAGTACTGCTCAGATG 451
Db |||||
QY 157 GGGGAATTGGCTGAGATGCCCATGTAGTGTCTCAGGCATGGAGAGTACTGCTCAGATG 98
Db |||||
QY 452 CCATTGTTGTGTGATCGAGAGAAACGAGAATGCTAGTCTAATAATACCCCTCCGTA 511
Db |||||
QY 97 CCATTGTTGTGTGATCGAGAGAAACGAGAATGCTAGTCTAATAATACCCCTCCGTA 38
Db |||||
QY 512 TG---CTAACCACTATTATTAATGGCCACATTTTC 545
Db |||||
QY 37 TGATCTAACCACTATTATTAATGGCCACATTTTC 1

RESULT 4
CA203163
LOCUS SCRFLL028D06.g FL1 Saccharum officinarum cDNA clone SCRFLL028D06
DEFINITION 5', mRNA sequence.
ACCESSION CA203163
VERSION CA203163.1 GI:35239153
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum complex.
REFERENCE 1 (bases 1 to 674)
AUTHORS Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
TITLE The libraries that made SUCST
JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at: <http://www.bcccenter.fcav.unesp.br>
Plate: 028 row: D column: 06
Seq primer: T7 Promoter Primer.

FEATURES
source
1..674
/organism="Saccharum officinarum"
/mol_type="mRNA"

/db_xref="taxon:4547"
/clone="SCRFLL028D06"
/lab_host="DH10B"
/clone_lib="FL1"
/note="Organ: Inflorescence at beginning of development (lcm-long); Vector: pSport1; Site_1: SalI; Site_2: NotI; An unidirectional cDNA library generated from [Inflorescence at beginning of development (lcm-long)]. cDNA was prepared from polyA+ mRNA using SuperScript plasmid system kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://sucust.lad.ic.unicamp.br/public>"

ORIGIN

Query Match 5.5%; Score 164.6; DB 6; Length 674;
Best Local Similarity 84.5%; Pred. No. 2.7e-33;
Matches 185; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 1453 GTTCTAAGGCCAGTCTCAGTGGGGTTTCATCAGAGTTTCATGGACATTAATAGCTGA 1512
Db |||||
QY 258 GTGCTAAGGCCAGTCTCAGTGGGGTTTCATCAGAGTTTCATGGACATTAATAGCTGA 199
Db |||||
QY 1513 TGTGACACCGTATTGATGAAGAGAGAGATGATAAGAGTTTCATGCGAGTAGAGAGATTT 1572
Db |||||
QY 198 TGTAGCACTATTATTATGAAGAGAGAGATGATAAGAGTTTCATGCGAGTAGAGAGATTT 139
Db |||||
QY 1573 CATGGGATGAACCTCTTCTCAGTGGTTTCCAAATATAGATGCTTGGTAAGAGGGCCA 1632
Db |||||
QY 138 CATGGAGATAAACCTCTTTTGCACCTGTTTCCAAATTTGGATGCGCTGGAAACAGTGACA 79
Db |||||
QY 1633 TGAATCTCTAGTACACCTGACCTAAGATGATGACT 1671
Db |||||
QY 78 TGAATCTCTAGTACACCTGACCTAAGAGCAGGTACAGT 40
Db |||||

RESULT 5
CA128031
LOCUS SCAGLR2018H12.g LR2 Saccharum officinarum cDNA clone SCAGLR2018H12
DEFINITION 5', mRNA sequence.
ACCESSION CA128031
VERSION CA128031.1 GI:35008477
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum complex.
REFERENCE 1 (bases 1 to 563)
AUTHORS Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
TITLE The libraries that made SUCST
JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at: <http://www.bcccenter.fcav.unesp.br>
Plate: 018 row: H column: 12
Seq primer: T7 Promoter Primer.

FEATURES
source
1..563
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"

```

/clone="SCAGLR2018H12"
/lab_host="DH10B"
/clone_lib="LR2"
/notes="Organ: Leaf roll from field grown adult plants
(small insert library); Vector: pSport1; Site 1: SalI;
Site 2: NotI; An unidirectional cDNA library generated
from [leaf roll from field grown adult plants (small
insert library)]. cDNA was prepared from polyA+ mRNA
using Superscript Plasmid System Kit (Invitrogen). The
double-strand cDNAs were fractionated in a sepharose
CL-2B 40cm-columns and fragments sizing between 0.8 and
1.5 Kb were directionally cloned into the vector. Details
of each source of RNA and library construction can be
obtained at http://sucsest.lad.ic.unicamp.br/public"

```

ORIGIN

```

Query Match      5.4%; Score 161.6; DB 6; Length 563;
Best Local Similarity 82.0%; Pred. No. 1.7e-32;
Matches 200; Conservative 0; Mismatches 39; Indels 5; Gaps 1;

QY 277 AGAATGTTGATCTGGAGAGTTTCTGAGATGTGACAAACGGGAGGTGATATCAAG 336
DB 318 AGGAATGTTGATCTAGAGAGTTCTGTGAGATGTGACAAACAGGAGGTGATATCAAG 377
QY 337 ATTCTGGGTACCCGGGAGATCGGCTCCATGTAGTTCGTCAGGCAATGGGGGA 396
DB 378 ATTCTGGGTACCTCGGAGANACAGCTTCATGTAGTTCGTCAGGCAATGGGGGA 437
QY 397 ATTGGCTGAGATGCCCCCA-----TGTAGTCTGAGGCAATGGAGATCTGCTGAGATG 451
DB 438 AATTGCTGAGATCCCCCGAGAGTAGCATCGTCAGGCAATGGAGAGATTGCTGAGATT 497
QY 452 CCATTGTTGTTAGTCGAGAGAACGAGAGATGCTAGTCTAATATACCTTCCGTA 511
DB 498 CCAATTGTTGGTAGATTGGGAAACAGAGAAATGCCCTGTCTAAATAATACCTTCCGAT 557
QY 512 TGCT 515
DB 558 TCCT 561

```

RESULT 6

```

CAL154373
LOCUS SCVPRZ2042F06.g R22 Saccharum officinarum cDNA clone SCVPRZ2042F06
DEFINITION 664 bp mRNA linear EST 24-SEP-2003
5', mRNA sequence.
ACCESSION CAL154373
VERSION CAL154373.1 GI:35061389
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum

```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum complex.

REFERENCE

1 (bases 1 to 664)
Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.
The libraries that made SUCST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br

Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at <http://www.bcccenter.fcav.unesp.br>
Plate: 042 row: F column: 06
Seq primer: T7 Promoter Primer.
Location/Qualifiers

FEATURES

source

1. .664

```

/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCVPRZ2042F06"
/lab_host="DH10B"
/clone_lib="R22"
/notes="Organ: Shoot-root transition zone from young plants
(small insert library); Vector: pSport1; Site 1: SalI;
Site 2: NotI; An unidirectional cDNA library generated
from [shoot-root transition zone from young plants (small
insert library)]. cDNA was prepared from polyA+ mRNA
using Superscript Plasmid System Kit (Invitrogen). The
double-strand cDNAs were fractionated in a sepharose
CL-2B 40cm-columns and fragments sizing between 0.8 and
1.5 Kb were directionally cloned into the vector. Details
of each source of RNA and library construction can be
obtained at http://sucsest.lad.ic.unicamp.br/public"

```

ORIGIN

```

Query Match      5.2%; Score 158; DB 6; Length 664;
Best Local Similarity 79.0%; Pred. No. 1.7e-31;
Matches 188; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 1436 TTTCGGGTAGGAGTAGGTTCTTAAGGCCAGTCTCAGTGGGGTTTCATCAGAGTTTCATG 1495
DB 159 TGTGTGCACAGTATGTTGTTAGTTCAGTCTTAATGGGGTTTGCATCAGAGTTTCATG 218
QY 1496 GACATTAAATAGCTGATGTGACACCGTATTGATGAAGAGAGATGATGAAGTTTCAT 1555
DB 219 GGCATTAAATATTGTTGATGTGACACCATTAATGAAGAGAGATGATGAAGTTTCAT 278
QY 1556 GCGAGTAGAGAGATTCATCGGGATGAACTCTTCTCACTGTTCCAAATATAGATG 1615
DB 279 GGGAGTAGAGAGATTCATGAGATGAGACTCTTATGCACTGTTTCCAAATCTGGATG 338
QY 1616 CATTGTTAAGAGGGCCATGAAATCTCTAGTGACACTGACCTTAAGATGAGATTGACTCT 1673
DB 339 TGTTGAAACAGTGATATGAATCCCACTGAAACTGCCCTTATGCGATGTTTCTTT 396

```

RESULT 7

```

CAL126310/c
LOCUS SCGLR1045G12.g LR1 Saccharum officinarum cDNA clone SCGLR1045G12
DEFINITION 694 bp mRNA linear EST 24-SEP-2003
5', mRNA sequence.
ACCESSION CAL126310
VERSION CAL126310.1 GI:35005077
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum

```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum complex.

REFERENCE

1 (bases 1 to 694)
Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.
The libraries that made SUCST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br

Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at <http://www.bcccenter.fcav.unesp.br>
Plate: 045 row: G column: 12
Seq primer: T7 Promoter Primer.
Location/Qualifiers

FEATURES

source

1. .694

/organism="Saccharum officinarum"

```

/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCSGLR1045G12"
/lab_host="DH10B"
/clone_lib="LRI"
/notes="Organ: Leaf roll from field grown adult plants
(large insert library); Vector: pSport1; Site_1: SalI;
Site_2: NotI; An unidirectional cDNA library generated
from [Leaf roll from field grown adult plants (large
insert library)]. cDNA was prepared from polyA+ mRNA
using SuperScript Plasmid System Kit (Invitrogen). The
double-strand cDNAs were fractionated in a sepharose
CL-2B 40cm-columns and fragments sizing between 0.8 and
1.5 Kb were directionally cloned into the vector. Details
of each source of RNA and library construction can be
obtained at http://sucet.lad.ic.unicamp.br/public"

```

ORIGIN

```

Query Match      5.2%; Score 157.8; DB 6; Length 694;
Best Local Similarity 81.3%; Pred. No. 2e-31;
Matches 183; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 1451 AGTTTCTAAGCCAGTCTCAGTGGGTTTCATCAGAGTTTCATCGACATTAAATAAGCT 1510
DB 373 AGTTTGGTAAGGCTAGTCTCAATGGAGTTTCATCAGAGTTTATAGACATTAAATATGCT 314

QY 1511 GATGTGACACCGTATTGATGAAGAGAGATGATGAAGAGTTTCATGCGAGTAGAGAGT 1570
DB 313 AATGTGGCACCGTATTATGAAGAGAGAGATGATAAAGTTTATGAGAGTAGAGATCGT 254

QY 1571 TTCATGGGAGTAAACTCTTCTTCACTGTTTCCAAAATATAGATCGATTTGTAAGAGGCG 1630
DB 253 TTCATGGAGATGAACCTCTCTGCACTGTTTCCAAAATATAGATCGTTGAAACCTAGA 194

QY 1631 CATGAAATCTCTAGTGACACGTACCTAAGATGAGATTGACTCTAG 1675
DB 193 CATGAAACTCCAACTGAAACTGGCCCTAAAGCAGTTGGCAACTTG 149

```

RESULT 8

```

CAL18511
LOCUS SCBGLR1099C01.g LRI Saccharum officinarum cDNA clone SCBGLR1099C01
5', mRNA sequence.
ACCESSION CAL18511
VERSION CAL18511.1 GI:34971819
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
complex.

```

REFERENCE

```

1 (bases 1 to 573)
Vettore,A.L., da Silva,P.R., Kemper,E.L. and Arruda,P.
The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenhariaia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.uneb.br
Plate: 099 row: C column: 01
Seq primer: P7 Promoter Primer.
Location/Qualifiers
1..573
/mol_type="mRNA"
/organism="Saccharum officinarum"

```

FEATURES

source

```

/db_xref="taxon:4547"
/clone="SCBGLR1099C01"
/lab_host="DH10B"
/clone_lib="LRI"
/notes="Organ: Leaf roll from field grown adult plants
(large insert library); Vector: pSport1; Site_1: SalI;
Site_2: NotI; An unidirectional cDNA library generated
from [Leaf roll from field grown adult plants (large
insert library)]. cDNA was prepared from polyA+ mRNA
using SuperScript Plasmid System Kit (Invitrogen). The
double-strand cDNAs were fractionated in a sepharose
CL-2B 40cm-columns and fragments sizing between 0.8 and
1.5 Kb were directionally cloned into the vector. Details
of each source of RNA and library construction can be
obtained at http://sucet.lad.ic.unicamp.br/public"

```

ORIGIN

```

Query Match      5.1%; Score 154.4; DB 6; Length 573;
Best Local Similarity 93.6%; Pred. No. 1.6e-30;
Matches 161; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 TCTAGACATAGGCATTGTAAAGGGTATGCCTCTCTTCAGTGCAGAAATTCATACCA 60
DB 402 TCTAGACATAGGCATTGTAAAGGGTATGCCTCTCTTCAGTGCAGAAATTCATACCA 461

QY 61 ACCTTAGGTATCTCTCTTCATAGAAATTTCTACCTGAGTAGGTTCCGTTCTGGTTGGAT 120
DB 462 ACCTTAGGTATCTCTCTTCATAGAAATTTCTACCTGAGTAGGTTCCGTTCTGGTTGGGA 521

QY 121 TTGTAGCGGGTTTCATGCAAAATAAGTTAGAAATCGTGCAAACTTGCAATGG 172
DB 522 TTGTAGCGGGTTTCATGCAAAATAAGTTAGAAATCGTGCAAACTTGCAATGG 573

```

RESULT 9

```

CWS12557/c
LOCUS CWS12557 634 bp DNA linear GSS 07-OCT-2004
DEFINITION Sorghum bicolor genomic clone 10510675, genomic survey sequence.
ACCESSION CWS12557.1 GI:53855339
VERSION GSS.
KEYWORDS Sorghum bicolor (sorghum)
SOURCE Sorghum bicolor
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 634)
Bedell,J.A., Budiman,M.A., Nunberg,A., Citek,R.W., Robbins,D.,
Jones,J., Flick,E., Rohlfling,T., Fries,J., Bradford,K.,
McMenamy,J., Smith,M., Holuman,H., Roe,B.A., Willey,G., Korfi,I.F.,
Rabinowicz,P.D., Lakey,N., McCombie,W.R., Jeddeloh,J.A. and
Martienssen,R.A.
Sorghum genome sequencing by methylation filtration
Unpublished (2004)
Contact: Bedell JA
Orion Genomics, LLC
4041 Forest Park Ave, St. Louis, MO 63108, USA
Tel: 314 615 6979
Fax: 314 615 5975
Email: jbedell@oriongenomics.com
Plate: 1 row: P column: 17
Seq primer: M13(40)Forward
Class: shotgun
High quality sequence stop: 634.
Location/Qualifiers
1..634
/organism="Sorghum bicolor"
/mol_type="genomic DNA"
/cultivar="Atx623"
/db_xref="taxon:4558"
/clone="10510675"
/clone_lib="Sorghum unfiltered library (LibID: 115)"

```

TITLE

JOURNAL COMMENT

FEATURES

source

prepared from purified nuclei was randomly sheared, end-repaired, size fractionated to enrich for the 0.5 to 5 kb fraction, ligated into HincII-digested pBCSK(-) vector and electroporated into *E. coli* cells. This is a methylation-filtered library."

ORIGIN

```

Query Match      4.9%; Score 148.6; DB 9; Length 717;
Best Local Similarity 79.9%; Pred. No. 6.7e-29;
Matches 175; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 1454 TTCTAAGCCAGTCTCAGTGGGGTTTCATCAGAGTTTCATGAGCATTAAATAAGCTGAT 1513
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 123 TFCATAAGACCAAGTCTCAATGGGGTTTCATTAGAGTTTCATGCACATTAATATCTGAT 182

QY 1514 GTGACACCGTATTGATGAAGAGAGAGATGATGAAGAGTTTCATGCCAGTAGAGAGTTTC 1573
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 183 GTGGCGCTATATTATGAAGAGAGAGATGATGAAGAGTTTCATGGGAGTAGAGAGTTTC 242

QY 1574 ATGGGGATGAAACTCTTCTTCACTGTTCCTCAAAATATAGATGCATTTGGTAAAGAGGCCAT 1633
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 243 ATCCCCATGAAGTCTCTATGCACCTGTTTCCAAATATTGATGTGTGGAACTGTGTAT 302

QY 1634 GAAATCTCTAGTGACACGACCTAAGATGAGATTGACTC 1672
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 303 GAAACTCCGTTGAGGAGTGGCCTAAGAGTAGAGAGTTTC 341

```

RESULT 12

```

CA220522
LOCUS SCRUF4023E04.g FL4 Saccharum officinarum cDNA clone SCRUF4023E04
DEFINITION 5', mRNA sequence.
ACCESSION CA220522
VERSION CA220522.1 GI:35273911
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
complex.
REFERENCE 1 (bases 1 to 669)
AUTHORS Vettore,A.L., da Silva,P.R., Kemper,E.L. and Arruda,P.
TITLE The libraries that made SUCEST
JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br

```

Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at <http://www.bcccenter.fcav.unesp.br>
 Plate: 023 row: E column: 04
 Seq primer: T7 Promoter Primer.

FEATURES

source

```

1..669
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCRUF4023E04"
/lab_host="DH10B"
/clone_lib="FL4"
/notes="Organ: Developed inflorescence and rachis (20cm-long); Vector: pSport1; Site 1: SalI; Site 2: NotI; An unidirectional cDNA library generated from [Developed] inflorescence and rachis (20cm-long)]. cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing

```

ORIGIN

```

Query Match      4.9%; Score 147.8; DB 6; Length 669;
Best Local Similarity 86.2%; Pred. No. 1.1e-28;
Matches 175; Conservative 0; Mismatches 27; Indels 1; Gaps 1;

QY 1458 TAAGCCAGTCTCAGTGGGGTTTCATCAGAGTTTCATGAGCATTAAATAAGCTGATGTA 1517
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 427 TAAGCCAGTCTCAAT-GAGTTTCATCAGAGTTTCATAGACATTAATATGCTGATGG 485

QY 1518 CACCGTATTGATGAAGAGAGAGATGATGAAGAGTTTCATGCCAGTAGAGAGTTTCATGG 1577
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 486 TACATATTATGAAGAGAGAGATGATGAAGAGTTTCATGGAGTAGAGAGTTTCATGG 545

QY 1578 GGATGAAACTCTTCTTCACTGTTCCTCAAAATATAGATGCATTTGGTAAAGAGGCCATGAAA 1637
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 546 GGATGAAACTCTTCTGCACTGTTCCTCAAAATCTGATGTGTGGAAACAGTGACATGAAA 605

QY 1638 TCTCTAGTGACACTGACCTAAGA 1660
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 606 TTCCCACTGAGACTGGCCTAATA 628

```

RESULT 13

```

CL191874
LOCUS CL191874
DEFINITION 740 bp DNA linear GSS 06-JAN-2004
104 414 10939851 114 32247 012 Sorghum methylation-filtered library (LibID: 104) Sorghum bicolor genomic clone 10939851, genomic survey sequence.
ACCESSION CL191874
VERSION CL191874.1 GI:40704397
KEYWORDS GSS.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 740)
AUTHORS Budiman,M.A., Flick,E., Jones,J., Nunberg,A., Citek,R.W., Robbins,D., Rohlfing,T., Bradford,K., Fries,J., McMenamy,J., Trani,L., Isak,A., Zimmerman,C., Lakey,N. and Bedell,J.A.
TITLE GeneThresher methylation filtered genomic sequences from Sorghum bicolor
JOURNAL Unpublished (2004)
COMMENT Contact: Bedell JA
Orion Genomics, LLC
4041 Forest Park Ave, St. Louis, MO 63108, USA
Tel: 314 615 6979
Fax: 314 615 5975
Email: jbedell@oriongenomics.com
Plate: 414 row: f column: 03
Seq primer: M13/pUC Forward
Class: shotgun
High quality sequence stop: 740.

```

FEATURES

source

```

1..740
/organism="Sorghum bicolor"
/mol_type="genomic DNA"
/cultivar="ATx623"
/db_xref="taxon:4558"
/clone="10939851"
/clone_lib="Sorghum methylation-filtered library (LibID: 104)"
/notes="Organ: leaf; Vector: pBCSK(-); Site 1: HincII; DNA prepared from purified nuclei was randomly sheared, end-repaired, size fractionated to enrich for the 0.5 to 5 kb fraction, ligated into HincII-digested pBCSK(-) vector and electroporated into E. coli cells. This is a methylation-filtered library."

```

between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://seucent.lad.ic.unicamp.br/public>

ORIGIN

Query Match 4.9%; Score 147.4; DB 9; Length 740;
 Best Local Similarity 74.2%; Pred. No. 1.4e-28;
 Matches 241; Conservative 0; Mismatches 76; Indels 8; Gaps 4;

QY 1340 ACATTAGTTGGTATCATTAATATTATATAT--TATCATATAAATTTGATCAACTTGAG 1397
 |||||
 Db 220 ACCTTTAGTTGATATCATATAAATATTATATCTACCATATATAATTTGGTTAAACTTAAG 279
 |||||

QY 1398 ATGCTTTGACTCTTCAAGATCTTGGAAATCTTGAATCACTTATATCTTGGGGTAGGAGTGGTTTC 1457
 |||||
 Db 280 ATGCTTTAACTATCTAAGATTTTGAATGACTTTGTCAITTTGGACGGAGGAGTAATCC 339
 |||||

QY 1458 TAAGGCCAGTCTCAGTGGGGTTTCATCAGAGTTTTCATGACATTAATAAGCTGATGGA 1517
 |||||
 Db 340 TTAGGCCAGTCTCAATGGGATTTTCAITAGAGTTTTCATGCACATTAATAATATGC--ATATGA 397
 |||||

QY 1518 CACCGTATTGATGAAGAGAGAGATGATAAGAGTTTTCATCGGACTAGAGAGTTTTCATGG 1577
 |||||
 Db 398 CACTGATTTATGA--AGAGAGATGATAGAAATTTCAATAAGTAGAAATAATTTTATCC 455
 |||||

QY 1578 GGATGAACACTCTTCTTCACTGTTTCCAAAA--TATAGATGCATTTGGTAAGAGGGCCCATGA 1635
 |||||
 Db 456 CTATAAACTCTTTTGGCTATTTTAAATCTAATAATGTTGTAGAACTAGATCATCA 515
 |||||

QY 1636 AATCTCTAGTGAAGTCACTGACCTAAGA 1660
 |||||
 Db 516 AATTACGATTGATGATGGCCTTAAA 540
 |||||

RESULT 14

BZ693362
 LOCUS BZ693362 1273 bp DNA linear GSS 02-JUL-2003
 DEFINITION SP_Ba0033E12.rp_Sp_Ba Sorghum propinquum genomic clone
 SP_Ba0033E12 3', genomic survey sequence.
 ACCESSION BZ693362
 VERSION BZ693362.1 GI:28385166
 KEYWORDS GSS.
 SOURCE Sorghum propinquum
 ORGANISM Sorghum propinquum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Sorghum.
 1 (bases 1 to 1273)
 Rabinowitz, P.D., O'Shaughnessy, A.L., Balijs, V., Dedhia, N.,
 Katzenburger, F., King, L., Miller, B., Muller, S., Nascimento, L.,
 Zutavern, T., Palmer, L., McCombie, W.R., and Martienssen, R.A.
 Genomic shotgun sequences from Sorghum bicolor (methyl-filtered)
 Unpublished (2002)
 Contact: W. Richard McCombie
 Lita Amentberg Hazen Genome Sequencing Center
 PO Box 100, Cold Spring Harbor, NY 11724, USA
 Tel: 516 367 8884
 Fax: 516 367 8874
 Email: mcombie@cshl.org
 Plate: ic83 row: e column: 09
 Seq primer: -21M13UnivRev
 Class: shotgun
 High quality sequence stop: 787.
 Location/Qualifiers
 1..787
 /organism="Sorghum propinquum"
 /mol_type="genomic DNA"
 /db_xref="taxon:132711"
 /clone="Sp_Ba0033E12"
 /clone_lib="SP_Ba"
 /note="Vector: pBelobAC11; Site_1: HindIII; Site_2: HindIII; Paterson lab BAC library (HindIII)"

FEATURES

source
 1..1273
 /organism="Sorghum propinquum"
 /mol_type="genomic DNA"
 /db_xref="taxon:132711"
 /clone="Sp_Ba0033E12"
 /clone_lib="SP_Ba"
 /note="Vector: pBelobAC11; Site_1: HindIII; Site_2: HindIII; Paterson lab BAC library (HindIII)"

ORIGIN

Query Match 4.7%; Score 141.8; DB 8; Length 1273;
 Best Local Similarity 82.2%; Pred. No. 5.8e-27;
 Matches 175; Conservative 0; Mismatches 37; Indels 1; Gaps 1;

QY 1455 TTCTAAGCCAGTCTCAGTGGGGTTTTCATCAGAGTTTTCATGGACATTAATAAGCTGATG 1514
 |||||
 Db 130 TACTAAGGCCAGTCTCAAT--GGGTTTTCATTAGAGTTTTCATGGATATTAATATGCTGATG 188
 |||||

QY 1515 TCACACCGTATTGATGAAGAGAGAGATGATAGAGTTTTCATCGAGTAGAGAGATGATCA 1574
 |||||
 Db 189 TGGCACCGTATTAAATGAATAAGAGGTGATAGAGTTTTCATGAGAGTAGTAGAGATTTA 248
 |||||

QY 1575 TGGGGATGAAACTCTTCTTCACTGTTTCCAAATATAGATGCATTTGGTAAGAGGGCCCATG 1634
 |||||
 Db 249 TAGGGATGAACTCTTGTACATTTGTTTTTAAATATGATGTTGTGGAACTGACCCATG 308
 |||||

QY 1635 AATCTCTAGTGAAGTCACTGACCTAAGATGATTT 1667
 |||||
 Db 309 AAACCTACACTGAGAATGATCTAAGAGTCTATT 341
 |||||

RESULT 15

BZ342445
 LOCUS BZ342445 787 bp DNA linear GSS 06-NOV-2002
 DEFINITION ic83e09.g1 WGS-SbicolorF (JM107 adapted methyl filtered) Sorghum
 bicolor genomic clone ic83e09 5', genomic survey sequence.
 ACCESSION BZ342445
 VERSION BZ342445.1 GI:24743123
 KEYWORDS GSS.
 SOURCE Sorghum bicolor (sorghum)
 ORGANISM Sorghum bicolor
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Sorghum.
 1 (bases 1 to 787)
 Rabinowitz, P.D., O'Shaughnessy, A.L., Balijs, V., Dedhia, N.,
 Katzenburger, F., King, L., Miller, B., Muller, S., Nascimento, L.,
 Zutavern, T., Palmer, L., McCombie, W.R., and Martienssen, R.A.
 Genomic shotgun sequences from Sorghum bicolor (methyl-filtered)
 Unpublished (2002)
 Contact: W. Richard McCombie
 Lita Amentberg Hazen Genome Sequencing Center
 PO Box 100, Cold Spring Harbor, NY 11724, USA
 Tel: 516 367 8884
 Fax: 516 367 8874
 Email: mcombie@cshl.org
 Plate: ic83 row: e column: 09
 Seq primer: -21M13UnivRev
 Class: shotgun
 High quality sequence stop: 787.
 Location/Qualifiers
 1..787
 /organism="Sorghum bicolor"
 /mol_type="genomic DNA"
 /db_xref="taxon:4558"
 /clone="ic83e09"
 /lab_host="JM107 or DH5a"
 /clone_lib="WGS-SbicolorF (JM107 adapted methyl filtered)"
 /note="Site_1: Xba I; Site_2: Xba I; The vector was filled in
 digested with XbaI and one nucleotide was added by fill in
 in the recessive 3' end. The genomic DNA was nebulized,
 end repaired, adaptor ligated and size fractionated using
 sephadex. The resulting fragments were between 0.8 and 3
 kb and were cloned into the vector (x/y reads in M13mp19,
 b/g reads in pUC19). The same ligation was transformed in
 either JM107 or DH5a."

FEATURES

source
 1..787
 /organism="Sorghum bicolor"
 /mol_type="genomic DNA"
 /db_xref="taxon:4558"
 /clone="ic83e09"
 /lab_host="JM107 or DH5a"
 /clone_lib="WGS-SbicolorF (JM107 adapted methyl filtered)"
 /note="Site_1: Xba I; Site_2: Xba I; The vector was filled in
 digested with XbaI and one nucleotide was added by fill in
 in the recessive 3' end. The genomic DNA was nebulized,
 end repaired, adaptor ligated and size fractionated using
 sephadex. The resulting fragments were between 0.8 and 3
 kb and were cloned into the vector (x/y reads in M13mp19,
 b/g reads in pUC19). The same ligation was transformed in
 either JM107 or DH5a."

ORIGIN

Query Match 4.6%; Score 139.4; DB 8; Length 787;
 Best Local Similarity 82.3%; Pred. No. 2.3e-26;

THIS PAGE BLANK (USPTO)